

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 07:32:31 ; Search time 10171 Seconds
(without alignments)
12680.637 Million cell updates/sec

Title: US-10-645-655-1

Perfect score: 4319
Sequence: 1 TCATAGTCGTTTACTAGT.....ACCTTTTATTCATATAT 4319

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	89.2	2.1	1202	28	BH854408 303 Genom
C 2	58.8	1.4	1101	29	AL063921 Drosophi1
C 3	58.4	1.4	994	13	BX414650
C 4	56.6	1.3	885	13	BX425603

C 5	56.2	1.3	712	13	BX416727	BX416727
C 6	55.6	1.2	1124	13	BX36282	BX36282
C 7	53.4	1.2	900	28	AZ531096	AZ531096
C 8	53.2	1.2	999	23	BX380865	BX380865
C 9	52.4	1.2	968	29	CNS06PBX	CNS06PBX
C 10	52.4	1.2	1201	13	BX461310	BX461310
C 11	51.2	1.2	926	14	CB986430	CB986430
C 12	50.8	1.2	617	14	CA743799	CA743799
C 13	50.2	1.2	770	28	BZ231524	BZ231524
C 14	50	1.2	1101	29	CNS017Y	CNS017Y
C 15	50	1.2	1111	13	BX376338	BX376338
C 16	49.6	1.1	872	28	AZ530255	AZ530255
C 17	49.6	1.1	1201	13	BX376337	BX376337
C 18	49.4	1.1	869	13	BX401089	BX401089
C 19	49.4	1.1	1164	13	BX425839	BX425839
C 20	49.2	1.1	1055	13	B087453	B087453
C 21	49	1.1	753	13	BX443342	BX443342
C 22	49	1.1	919	29	CNS05SR	CNS05SR
C 23	48.8	1.1	475	12	BM029212	BM029212
C 24	48.8	1.1	977	13	BX397883	BX397883
C 25	48.6	1.1	892	28	BZ256924	BZ256924
C 26	48.4	1.1	711	29	CNS011V8	CNS011V8
C 27	48.4	1.1	828	29	CNS011TX	CNS011TX
C 28	48.4	1.1	1201	29	CNS016E1	CNS016E1
C 29	48.4	1.1	1306	13	B0675354	B0675354
C 30	48.2	1.1	633	13	BX403519	BX403519
C 31	48	1.1	1180	13	BX436369	BX436369
C 32	47.8	1.1	710	12	BM163311	BM163311
C 33	47.8	1.1	988	14	CD249057	CD249057
C 34	47.8	1.1	1005	14	CB207771	CB207771
C 35	47.8	1.1	1201	13	BX356851	BX356851
C 36	47.8	1.1	1201	13	BX46717	BX46717
C 37	47.6	1.1	545	29	CE364861	CE364861
C 38	47.6	1.1	558	29	CE13657	CE13657
C 39	47.6	1.1	728	29	CE697734	CE697734
C 40	47.6	1.1	1016	29	CNS016J9	CNS016J9
C 41	47.4	1.1	849	29	CNS03664	CNS03664
C 42	47.2	1.1	1099	13	BX456575	BX456575
C 43	47.2	1.1	1101	29	CNS0182P	CNS0182P
C 44	47	1.1	1101	29	CNS006MC	CNS006MC
C 45	46.8	1.1	1079	13	BX398967	BX398967

ALIGNMENTS

RESULT 1
LOCUS BH854408/C 1202 bp DNA linear GSS 25-JUN-2002
DEFINITION 303 Genomic fragments from E. coli EC317 absent from B. coli K-12
LOCUS MG1655 Escherichia coli genomic 5', genomic survey sequence.

ACCESSION BH854408
VERSION BH854408.1 GI:21554464
KEYWORDS GSS.
SOURCE Escherichia coli
ORGANISM Escherichia coli

REFERENCE 1 (bases 1 to 1202)
AUTHORS Stocki S.L., Babluk L.A., Rawlyk N.L., Porter, A.A. and Allan B.C.
TITLE Identification of Genomic Differences between E. coli strains
Subtractive Hybridization and A. coli K-12 MG1655 using Suppression

JOURNAL Unpublished (2002)
COMMENT Contact: Stacy Stocki
Veterinary Infectious Disease Organization
University of Saskatchewan
120 Veterinary Road, Saskatoon, Saskatchewan, Canada, S7N 5E3
Tel: 3069667467
Fax: 3069667478
Email: stocki@duke.usask.ca
Fragment 303
PCR Primers

1. -1202

CRIGIN

Matches 20

55

REFERENCE

COMMENT

FEATURES

ORIGIN

Matches 11

15

Oy	1140	AATGCGAAGCCATCTATTGATTCATAAAACGAAGCATTAATCTTGCAATC	1195
Dd	442	TTTTTTTWTMTTMAATTTATTTTITTTTAAAAAAAAMMMWMTTATATKCCCCCYC	387
 RESULT 3 BX141650/c			
LOCUS	BX141650	994 bp	mRNA linear EST 15-MAY-2003
DEFINITION	BX141650 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCP001Y02		
ACCESSION	BX141650		
VERSION	BX141650.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Li, W.B., Gruber, C., Jessee, J. and Polayes, D. (bases 1 to 994) Full-length cDNA libraries and normalization Unpublished (2001)		
AUTHORS	Contact: Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segre@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6015.f Contact : Feng liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CSOCP001DG01NP1. Location/Qualifiers		
TITLE	1. .994		
JOURNAL	/organism="Homo sapiens"		
COMMENT	/mol_type="mRNA" /db_xref="taxon:9606" /clone="CSOCP001YN02" /tissue type="THYMUS" /clone_lib="Homo sapiens THYMUS" /note="vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. library was not normalized."		
 FEATURES source			
 ORIGIN			
Query Match	1.4%; Score 58.4; DB 13; Length 994;		
Best Local Similarity	21.2%; Pred. No. 0.18;		
Matches 136,	Conservative 214; Mismatches 291; Indels 2; Gaps 1;		
Oy	824	AAATGCTTAATTAAGGGATTAATCGGGAGAACCCTTTGAAGCAAAAGAAATG	883
Dd	968	WDDDDSKKDWMWAADAKMRAARDKMWAGARRRRRBDADWMAARFMDRTMWAAAR	929
Oy	884	GTTTCATATGTTGGCAATCTTATTTGTGATGAATTTTCGAAAGATTTACATACATC	943
Dd	928	KKAAADTTTKTKKQDAADMGDDMDMTAAWMAADWMAARAAATATWTDMDTAWW	869
Oy	944	ACTTACACCCGAGCTGTGATGAGTGCACAAATTGAGGAAATGATTAATGCTAGGG	1003
Dd	868	TRDXKATWADDADDDDKTKT--TTTWTMKMTDMVGATWGWADATWADAADAAR	811
Oy	1004	GTCATTAATCTCGAAATCGGAATCCATGAGAAATTAATTTAGCAATATGAG	1063
Dd	810	AWMAAAMWRPAAAMWAADADMDWMMWMDAADAAAMAAMDYTKDRAAMPADWAPAAAA	751
Oy	1064	TTTACCTTTGAAAGAGAGGATTAATGTCATATCCATGATGACGAGCAATATATTTA	1123
Dd	750	AAMDWGRGBADAGDKARAAMWDCAGGGRDXDRBMAARMDDADBAARPAAAAAATAMW	691
Oy	1124	TTCTCCAGCTTTAAACAATGAGAAACGCTATATTTTATGATCAAAAACAGATATTT	1183
Dd	690	ATWAKWKCDWMAAAAANDTAKRGKRKDADDDWMAWTTTTTTTTTAAAADKMKMKWT	631

[illegible]

```

QY 2744 TTATTTGGCTATTAAGCATTAATTAATTCACATGACGAGCGCATACGT 2803
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 686 KXKXGKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKX 627
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2804 ATTATCTGTTGCAACAGCAGCAAGAAACCCGAAACCTTGACAAATTAATTGGTGA 2863
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 626 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 567
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2864 AAGCAAGATATCAACGTTATCAGATTAAGCTCAATTTACTTTAGAAATGACACAGT 2923
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 566 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 507
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2924 TGATGAGCTGCAATTAAGTTAATTAAGTGAAGAAATGATGCGAATTCGCTTGATTA 2983
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 506 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 447
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2984 CCCATTAAGAGAGAGAAATTCACATGATTTAGTAAGCAGAGACAGACAGAA 3043
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 446 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 387
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3044 ATTAGAGCCCAACAGTTGACGAGCTGCTAATAACACAAACAGTGAAGCAAGTGG 3103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 386 KTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3104 GTCAAGAGAGAGCAGCAGCAGCGCTTCTGATACCTGCTGATCAAGCCTGTTAA 3163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 326 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3164 CGCATTAAGAGCCCAACAGCTGACTGACTGCTGAAACCAAAAAAGTAGGCAAAAAC 3223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 TNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3224 AAAAAAGTGGCTCAAAAAGA 3245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 206 AAAAAAAGTGGCTCAAAAAGA 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
BX416727/c 712 bp mRNA linear EST 15-MAY-2003
LOCUS BX416727 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION CS0DA011Y114 5-PRIME, mRNA sequence.
ACCESSION BX416727
VERSION BX416727.1 GI:30765629
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 712)
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@life.techn.com URL:
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue Genoscope sequence ID: CS0DA011B070P1.
Location/Qualifiers
1..712
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA011Y114"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/Note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector."

```

```

ORIGIN Library was not normalized.
Query Match 13%; Score 56.2; DB 13; Length 712;
Best Local Similarity 15.8%; Pred. No. 0.51;
Matches 71; Conservative 177; Mismatches 201; Indels 0; Gaps 0;

QY 3030 CAAGCAGAAAGCACTTGAAGACCAACAGTTGACCGACTCTAAAAACACAAACAGT 3089
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 712 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 653
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3090 GAGCCAAAAGTGGCTGCAAGAGAGCAGCAGAGAGCGTTCTGATACCTGCTGAT 3149
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 632 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 593
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3150 CAAAGCTGTTAAAGCATTAAGAGCAACAAAGTGAACGACTGCTGAAAACAAAAA 3209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 592 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 533
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3210 AGTAAGGCAAAAACAAAAGAGTGGCTGCAAAAAGACGCTGTTTCTGATCCCTGCTT 3269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 532 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3270 GATCAAGCCTGTCGATTAAGAGCGCAGCTGAGTTAGTATGATGCCCAACAGATCG 3329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 472 CAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3330 GAAAAAGCTGCTGCTCAAGAGAGGAAAAACAGCAGCAACAAAGACTTATC 3389
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 412 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3390 AGCGTTATTCAAATGAGGTTATGCAATATCTGCAACAGTAAATGATGCTTCT 3449
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 352 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3450 GTCAAGATGATTAAGTGCCTTTCTTGT 3478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 292 TTTTCTTTTCTTTTCTTTTCTTTTCTTTT 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
BX436282/c 1124 bp mRNA linear EST 15-MAY-2003
LOCUS BX436282 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP001Y001
DEFINITION BX436282 5-PRIME, mRNA sequence.
ACCESSION BX436282
VERSION BX436282.1 GI:30787521
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1124)
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@life.techn.com URL:
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue Genoscope sequence ID: CS0CAP001A010P1.
Location/Qualifiers
1..1124
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP001Y001"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/Note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

```



```

QY 3403 ATATGCGTTATGAGATTATCTGCAACAGTAATATGATGCTTCTTCTTCTTCAAGATGAT 3462
DB 134 AAGATGCAAAAGTGAAGAAACCAACACCAACAAATATAGAGATTAATGAAACCAACAA 135
QY 3463 TGAATGCTCTTTTGTAGATCAAGCAATCTGCTGTCGCAATATGTCACAGGATA 3522
DB 134 AAGAAAGATGAAAGAAATATCAACAAACCAAGAGATGAAATTAAGTAAAGACACAA 75
QY 3523 AAAGACGCTATGATCTGA 3541
DB 74 AAGCAACAGTACTCTGA 56

RESULT 3
EX380865 999 bp mRNA linear EST 08-MAY-2003
LOCUS EX380865
DEFINITION EX380865 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
VERSION EX380865
KEYWORDS EST.
ACCESSION EX380865.1 GI:30439622
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
AUTHORS Full-length cDNA libraries and normalization
TITLE Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91036 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3221.r. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1058BC05NP1&cluster=3221.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0D1058BC05NP1.

FEATURES
Source
1. 999
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1058YE10"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pOWS-PORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.2%, Score 53.2; DB 13; Length 999;
Best Local Similarity 24.6%; Pred. No. 2;
Matches 112; Conservative 118; Mismatches 225; Indels 0; Gaps 0;

QY 695 TGGAGCAGTAATGATATCGTATTTGGAGGCGCATGTTGCTAAAGCGGAGAAATGCG 754
DB 399 TGAASANSSITANNSTANNSCANASSSSGSSCNSSSTSTNNSGANNSSAVAAATSS 458
QY 755 TGCATTCAGCATTCAGGCTCAAGGGGACAGTGGTCTCCGATGTTATTTATGATGC 814
DB 459 TSSAANTNSAASINTTBTTABABANTATBTGNTTANNTTBABNTBBBAADPWATATA 518
QY 815 TGAATAAACAATAATGTTAATTAATGACATTAAGGAGAGGACCCCTTTGAAGGCA 874
DB 519 TAAATKKAADAAAANDGKARATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 578
QY 875 AGAATAATGSGTTTCAATGTTGCGCAACCTATATTTTATGAAATTTTGGAAAGATTT 934
DB 934

```

```

DB 579 WWAATAATTTTWTATTAKKNDKAAKAAADADAADAAAAAAXKXKXKXKXKXKXKXKX 638
QY 935 ACATACATCACTTTRACCCGACCTGTAATGAGCTACACATTAAGTGAATGATTA 994
DB 639 AADADAMADAKKKKDXADAGDAAKKKKRKKANKAAADNKKKAAAAAARAAAAAAXKX 698
QY 995 TGGTCAGGCTGTAATTAATCGAATACAGAAATCCATCAATTAATTAATTAATGATG 1054
DB 699 AAAAAKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKX 758
QY 1055 AATATAGATTTACCTTGAAGAGAGATTAAGTTTATATCTATATGACGAGCC 1114
DB 759 AAAAAANKDAAAADADAAAAAANDKRTATTAATTTTATTTTATTTTATTTTATTTT 818
QY 1115 TATATTTTCTCTCACTTTAAACATGAGAAA 1149
DB 819 AKADADDTKMAAAMAAKMAAAMATTTAAANAAA 853

RESULT 9
CNS06PBX 968 bp DNA linear GSS 05-JUL-2001
LOCUS CNS06PBX
DEFINITION T7 end of clone AV0AA013b12 of library AV0AA from strain CBS 379 of
Saccharomyces exiguus, genomic survey sequence.
ACCESSION AL409171
VERSION AL409171.1 GI:12176428
KEYWORDS GSS.
SOURCE Saccharomyces exiguus
ORGANISM Saccharomyces exiguus
REFERENCE Saccharomyces exiguus
AUTHORS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
1 (bases 1 to 968)
2 (bases 1 to 968)
3 (bases 1 to 968)
TITLE yeast species for molecular evolution studies
JOURNAL FEMS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
PUBMED 11152876
AUTHORS Bon, E., Neuveglise, C., Lepingle, A., Wincker, P., Artiguenave, F.,
Galliard, C. and Casaregola, S.
Genomic exploration of the hemiascomycetous yeasts: 6.
Saccharomyces exiguus
FEMS Lett. 487 (1), 42-46 (2000)
JOURNAL 20584716
MEDLINE 11152881
PUBMED 11152881
REFERENCE 3 (bases 1 to 968)
AUTHORS Genoscope.
TITLE Direct Submision
JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5736, 91057 EVRY cedex, FRANCE. (E-mail :
seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES
Source
1. 968
Location/Qualifiers
/organism="Saccharomyces exiguus"
/mol_type="genomic DNA"
/strain="CBS 379"
/db_xref="taxon:34358"

```

```

/misc_feature
      /clone="AV08A013H12"
      /clone_lib="AV08A"
      /note="end : 17*"
      /note="complement (<2. .>967)
      /note="similar to Saccharomyces cerevisiae ORF YLR087c {
      /note="hypothetical protein }
      /note="1 putative frameshift(s) *"
      /evidence="not_experimental"

```

ORIGIN

```

Query Match      1.2%; Score 52.4; DB 29; Length 968;
Best Local Similarity 39.7%; Pred. No. 2.9; Indels 0; Gaps 0;
Matches 155; Conservative 32; Mismatches 203;

QY 2860 TTGAAGCAAGATTAATCAACCGTTATCAGATAGCTCAATTTAGTGAAGTACC 2919
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 207 TWWRAAAAGATCAGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 266
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2920 ACGTTGATGAGGTGATTCATTTATTAATTAATTAATTAATTAATTAATTAAT 2979
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 267 CRRRAATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 326
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2980 ATAAACCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3039
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 327 TTAATAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 386
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3340 GAACATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3099
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 387 AAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 446
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3100 TGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3159
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 447 ATTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 506
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3160 TAAAGATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3219
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 507 TAAAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 566
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3220 AAACAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3249
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 567 AAACAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 596
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
BX461310      1201 bp mRNA linear EST 22-MAY-2003
LOCUS        BX461310 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION   CSDDF025YA16 5-PRIME, mRNA sequence.
ACCESSION    BX461310
VERSION      BX461310.1 GI:31031291
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 1201)
AUTHORS     Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 Evry cedex - France
              Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. Contact : Feng Liang Email : fliang@life.techn.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Paraday Avenue Genoscope sequence ID : CSDDF025YA160Pl.
FEATURES
  source
    1..1201
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CSDDF025YA16"

```

```

/tissue_type="FETAL BRAIN"
/dev stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dt) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

```

Query Match      1.2%; Score 52.4; DB 13; Length 1201;
Best Local Similarity 22.8%; Pred. No. 2.9;
Matches 104; Conservative 154; Mismatches 192; Indels 6; Gaps 1;

QY 8 TCGTTTAATGATTTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 67
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 721 TTKRRRAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 68 AACTGATTTGCTTATTTTATTAACCGTTGATTTTATTAAGAGAGAGAGAGAG 127
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 781 RRRGRGGGGGGGWWDDDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 840
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 128 GTGGGCTGCTACG-----CTTATTGGAGATTTATTAACATATTAATGCTGATTTTG 181
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 841 DGGGGKKKKKKDDAAAADDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 900
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 182 CGAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 241
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 901 KKKKKKKKKDDAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 GCAATTAATGTCACATCAATGACAAAGCCCGAGATTTTCTGATGATGACAG 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 961 AAAAABKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1020
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 TAACGCGTGGGACCTGCTGTAATTAATTAATTAATTAATTAATTAATTAAT 361
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1021 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1080
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 ATATACAGATTTGATTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1081 KAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 422 TAAGATTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 457
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1141 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1176
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 11
CB986430/c    926 bp mRNA linear EST 01-MAY-2003
LOCUS        CB986430/c
DEFINITION   AGENCOURT 13645938 NIH-MGC 184 Homo sapiens cDNA clone
              IMAGE:30329463 5', mRNA sequence.
ACCESSION    CB986430
VERSION      CB986430.1 GI:30280950
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 926)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgsbse@mail.nih.gov
              Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/ILNL at:
              http://image.llnl.gov
              Plate: NDCh141 row: 1 column: 16

```

High quality sequence stop: 142.
 FEATURES
 Location/Qualifiers

1..926
 /organism="Homo sapiens"
 /mol_type="rRNA"
 /db_xref="taxon:9606"
 /clone="TMAG:30329463"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1lb="NIH MGIC 184"
 /note="Organ: Pooled-glandular; Vector: pDR-LIB; Site_1: SfiI (ggccatagcc); Site_2: SfiI (ggccgcctggcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATATGCGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCCGGCGCGCATG-dT(30)BN-3' (where S = A, C, or G and N = A, C, G, or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH MGIC library."

ORIGIN

Query Match 1.2%; Score 51.2; DB 14; Length 926;
 Best Local Similarity 44.1%; Pred. No. 5.2;
 Matches 215; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

2758 AACGCGTAATTAATTAATTCATATGCGTGAAGGCGATTAATATATCTGTGCA 2817
 |||
 559 AA 500
 2818 ACACAGGCAAGAACCCGAAACCTTGAGCAATTAATCTTGTGAGCAAGATATC 2877
 |||
 499 ACTAAGCAACAACAAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 440
 2878 AACGTTATGATGATGCTCAATTTCTTGAATAATGACGCTGATGAGTGATG 2937
 |||
 439 AAGAGAGAAACAAAGACAAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 380
 2938 TACGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2997
 |||
 379 AAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 320
 2998 AGAATTCACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3057
 |||
 319 AA 260
 3058 AAGTTGACCGATCTTAACACAAAGCTGAGCCAAAGTGCCTCAAGAGAGCAG 3117
 |||
 259 GAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 200
 3118 CGAGAGCAGCGTTCTGATACCTGCTGATCAAGCGCTTTAAAGCATTAAGAGCA 3177
 |||
 139 AAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 140
 3178 AACAGCTGAATGACTGCTGAAACACAAAGAAAGTGAAGCAAAAAAGTGGCT 3237
 |||
 139 AAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 80
 3238 CAAAAAGA 3245
 |||
 79 AAAAAA 72

RESULT 12
 CA743799 617 bp mRNA linear EST 26-NOV-2002
 LOCUS
 DEFINITION wtls.pk005.p2 wtls Triticum aestivum cDNA clone wtls.pk005.p2 5'
 ACCESSION CA743799
 VERSION CA743799.1 GI:25559622
 KEYWORDS EST.

SOURCE
 ORGANISM
 Triticum aestivum (bread wheat)

Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poaceae; Triticaceae; Triticum.
 1 (bases 1 to 617)
 Tingley, S.V., Powell, W., Walters, P., Dolan, M., Hanney, C., Yuan, Z.,
 Mao, G., Caraher, N. and Hanafey, M.K.

TITLE
 JOURNAL
 Unpublished (2002)
 Contact: Scott V. Tingley
 Crop Genetics
 R. I. Dupont de Nemours and Company
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
 Tel: 302-631-2602
 Fax: 302-631-2607
 Email: Scott.V.Tingley@usa.dupont.com

FEATURES
 source

1..617
 Location/Qualifiers
 /organism="Triticum aestivum"
 /mol_type="rRNA"
 /db_xref="taxon:4565"
 /clone="wtls.pk005.p2"
 /tissue_type="leaf"
 /lab_host="DH10B"
 /clone_1lb="wtls"
 /note="Vector: pGEM-T Easy; Site 1: SmaI; R-band (susceptible) wheat leaves infected with Septoria tritici strain A, 24 hours after infection, subtracted w/ comparable uninfected leaves"

ORIGIN

Query Match 1.2%; Score 50.3; DB 14; Length 617;
 Best Local Similarity 46.2%; Pred. No. 6.4;
 Matches 163; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

3017 AGTAAGCAGAGCAGACAGACATTAAGAGCCAAACAGTTGAACCGACTGCTAA 3076
 |||
 161 AGAAG 220
 3077 AACCAACAGAGTACCAAGTCCGCTCAAGAGAGAGAGAGAGAGAGAGAGAG 3136
 |||
 221 AGAAG 280
 3137 TACCTGCTGATCAAGCGCTGTTAAAGCATTAAGAGCAACAGCTGAAGCTGC 3196
 |||
 281 AGAAG 340
 3197 TGAAGACAAAGAGTAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3256
 |||
 341 AGAAG 400
 3257 TGATCCCTGCTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3316
 |||
 401 AGAAG 460
 3317 CCCACAGCATCGGAG 3376
 |||
 461 AG 520
 3377 AAAAAA 3382
 |||
 521 AGAAG 526

RESULT 13
 B2231524/c 770 bp DNA linear GSS 12-OCT-2002
 LOCUS
 DEFINITION CH230-438K4.TV CHORI-230 Segment 2 Ratius norvegicus genomic clone
 ACCESSION B2231524
 VERSION B2231524.1 GI:23890065

KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 770)
Zhuo, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
Shwartsbeyn, A., Gebregiorgis, E., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
JOURNAL Unpublished (1999)
COMMENT Other GSSs: CH230-438K4.TU
Contact: Shaying Zhao
Department of Bukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@tigr.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ordering information.html). BAC end
page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 438 row: K column: 4
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..770
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SnHsd/MCW"
/db_xref="taxon:10116"
/clone_id="CH230-438K4"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: pTARHAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match 1.2%; Score 50.2; DB 28; Length 770;
Best Local Similarity 49.8%; Pred. No. 8.4; Indels 0; Gaps 0;
Matches 127; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
Db 2992 AAGCAGGAGATTGCACTATTAGTAAGACAGACAGACAGACATTAGAG 3051
|||||
712 AAG 653
3052 CCAACCAAGTTGACCGACTGCTAAACCAACCAAGCTGACCAAGCTCGCTAAGA 3111
|||||
652 CAG 593
3112 GAGCAGGAG 3171
|||||
592 AAG 533
3172 AAGCCAAACAGCTGACTGCTCTGTAACCAACCAACCAAGTAAAGCAACCAACCA 3231
|||||
532 AAG 473
3232 TCCGCTCAAAAAGAG 3246
|||||
Db 472 GAGAGAGAGAGAGAG 458

RESULT 14
LOCUS CNS017Y 1101 bp DNA linear GSS 26-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC

ACCESSION BACN37B08 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
VERSION A1108676
KEYWORDS A1108676.1 GI:5628980
SOURCE GSS.
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyritroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage;
BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.crs.fr
- Web: www.genoscope.crs.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billard at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_id="BACN37B08"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"

ORIGIN
Query Match 1.2%; Score 50; DB 29; Length 1101;
Best Local Similarity 30.9%; Pred. No. 9;
Matches 98; Conservative 83; Mismatches 132; Indels 4; Gaps 1;
Db 777 AAGGGGAGACGGTTCGATGTTTATTTATGATCGTAAACCAACCAATGTTAA 836
|||||
750 AATGAGATTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809
837 AATGAGATTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 896
|||||
810 AATGAGATTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 869
897 CGCAATCTTATTTGATGAAATTTGCAAGAGATTACATCACTTACACCCGA 956
|||||
870 AATGAGATTTATTTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929
957 GCTGTAATGAG 1012
|||||
930 AATGAGATTTATTTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 989
1013 TCAGAAATGAG 1072
|||||
990 AATGAGATTTATTTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1049
1073 GAAAG 1089
|||||
Db 1050 AATGAG 1066

RESULT 15
EX376338/c
LOCUS EX376338 1111 bp mRNA linear EST 08-MAY-2003
DEFINITION EX376338 Homo sapiens NEUROBLASTOMA COR 25-NORMALIZED Homo sapiens
CDNA clone CS00C026YF21 5-PRIME, mRNA sequence.
ACCESSION EX376338
VERSION EX376338.1 GI:30446522
KEYWORDS EST.
SOURCE Homo sapiens (human)

XX

PD 29-FEB-1996.
 XX 16-AUG-1995; 95MO-US010661.
 XX 25-AUG-1994; 94US-00296791.
 XX (UNIV) UNIV WASHINGTON.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX St Gene JW, Falkow S;
 XX WPI; 1996-151147/15.
 DR P-PSDB; AAR92768.
 XX Hemophilus adhesion and penetrator protein and corresponding DNA - used
 PT to produce vaccine against H. influenzae infection.
 XX Claim 5; Fig 6; 105bp; English.
 CC The sequence encodes a Haemophilus influenzae adhesion and penetration
 CC protein. The sequence (hap gene) includes putative -10 and -35 sequences
 CC and a putative rho-independent terminator 3' to the hap stop codon. The
 CC terminator contains interrupted inverted repeats, with the potential for
 CC forming a hairpin structure containing a loop of 3 bases and a stem of 8
 CC bases, followed by a stretch rich in T residues. The gene product is
 CC first synthesised as a preprotein, which is transported to the periplasm,
 CC followed by insertion of the C-terminal beta-domain into the outer
 CC membrane, possibly forming a pore, and export of the N-terminal fragment
 CC through the outer membrane, followed by autoproteolytic cleavage and
 CC secretion of the mature protease, leaving an outer membrane protein
 CC fragment. The gene may be inserted in a vector and expressed in
 CC recombinant host cells, for use as a recombinant vaccine. The gene
 CC product may also be used in diagnostic monoclonal antibody production.
 SQ Sequence 4319 BP; 1497 A; 776 C; 891 G; 1155 T; 0 U; 0 Other;
 Query Match 100.0%; Score 4319; DB 2; Length 4319;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCATAGTCTTTAATGATATTTTAAAGCAAAATTTCTTAATTAATTAACATTA 60
 DB 1 TCATAGTCTTTAATGATATTTTAAAGCAAAATTTCTTAATTAATTAACATTA 60
 QY 61 TGAATAAACTGATATGCTTTAATTTTAAAGCTTGATTTATAGGATATAT 120
 DB 61 TGAATAAACTGATATGCTTTAATTTTAAAGCTTGATTTATAGGATATAT 120
 QY 121 CGCAAGCGTGGCTGGTCACTATTTTGGATGATTAACATATTAATGATTTG 180
 DB 121 CGCAAGCGTGGCTGGTCACTATTTTGGATGATTAACATATTAATGATTTG 180
 QY 181 CCGAGATTAAGGAGAGTTCACAGTTGGGCTCAAAATTTAAGTTTATACCAAG 240
 DB 181 CCGAGATTAAGGAGAGTTCACAGTTGGGCTCAAAATTTAAGTTTATACCAAG 240
 QY 241 GGCATTTAGTGGCAATCAATGACAAAGCCCGCATGATTTTCTGATGTCAC 300
 DB 241 GGCATTTAGTGGCAATCAATGACAAAGCCCGCATGATTTTCTGATGTCAC 300
 QY 301 GTAAAGCGTGGCAGCTTGTGTAATAATATATTTGAGCGTGGCAATTAAG 360
 DB 301 GTAAAGCGTGGCAGCTTGTGTAATAATATATTTGAGCGTGGCAATTAAG 360
 QY 361 GATTAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420
 DB 361 GATTAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420
 QY 421 ATAAGATTTGTAAGCAATTAATACAAAGATTAATTTACATCTTTAGAGCAT 480
 DB 421 ATAAGATTTGTAAGCAATTAATACAAAGATTAATTTACATCTTTAGAGCAT 480
 QY 481 ACCATATTCAGATTAATTAATTTGTTACAGAAAGCGCTCCAAATGATGACTTGA 540

DB 481 ACCATATTCAGATTAATTAATTTGTTACAGAAAGCGCTCCAAATGATGACTTGA 540
 QY 541 ATATGATGGCAGTATCTTATTCAGATTAATTAATTCAGATGTTGTTGTTGTTG 600
 DB 541 ATATGATGGCAGTATCTTATTCAGATTAATTAATTCAGATGTTGTTGTTGTTG 600
 QY 601 CTGAGCGGATTTGGCGAAATGATCAAGCAAGCGCAAGTTCGGTGCATATC 660
 DB 601 CTGAGCGGATTTGGCGAAATGATCAAGCAAGCGCAAGTTCGGTGCATATC 660
 QY 661 ATTATCTGACGCTGGCAATACCAATCAGCTGGAGCAGTATGATATGATAT 720
 DB 661 ATTATCTGACGCTGGCAATACCAATCAGCTGGAGCAGTATGATATGATAT 720
 QY 721 TGGAGGCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 780
 DB 721 TGGAGGCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 780
 QY 781 GGGACGTTGTTCTCGATGTTTATTTATGATGCTGAAAGCAAAATGTTATTAATG 840
 DB 781 GGGACGTTGTTCTCGATGTTTATTTATGATGCTGAAAGCAAAATGTTATTAATG 840
 QY 841 GATATTAAGGAGGCAACCTTTTGAAGGCAAGAAATGGGTTTCAATTTGTTGCA 900
 DB 841 GATATTAAGGAGGCAACCTTTTGAAGGCAAGAAATGGGTTTCAATTTGTTGCA 900
 QY 901 AATCTTATTTGATGAAATTTTGAAGAGTTTACATACATCTTTACCCGAGCTG 960
 DB 901 AATCTTATTTGATGAAATTTTGAAGAGTTTACATACATCTTTACCCGAGCTG 960
 QY 961 GTATGAGTGTACCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1020
 DB 961 GTATGAGTGTACCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1020
 QY 1021 CAGGAATACCATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
 DB 1021 CAGGAATACCATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
 QY 1081 AGATTAAGTCTATATCTATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 1081 AGATTAAGTCTATATCTATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 1141 ATGAGAAACGCTATTTTATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1141 ATGAGAAACGCTATTTTATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1201 TTAACCAAGGCGGCTGCTTATTTTGAAGGTTAATTTTACATGATCTCAATTTCTA 1260
 DB 1201 TTAACCAAGGCGGCTGCTTATTTTGAAGGTTAATTTTACATGATCTCAATTTCTA 1260
 QY 1261 ACCAACTTGGCAGAGGCTGGGATATGATGATGATGATGATGATGATGATGATG 1320
 DB 1261 ACCAACTTGGCAGAGGCTGGGATATGATGATGATGATGATGATGATGATGATG 1320
 QY 1321 TAAATGGCTGGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 DB 1321 TAAATGGCTGGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 QY 1381 CCAAGGCGGAAATTAAGGTTGATGATGATGATGATGATGATGATGATGATGAT 1440
 DB 1381 CCAAGGCGGAAATTAAGGTTGATGATGATGATGATGATGATGATGATGATGAT 1440
 QY 1441 AGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 DB 1441 AGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 QY 1501 GAGCTGTTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 DB 1501 GAGCTGTTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 QY 1561 GTGCTGCTGCTGATCTTAACGCGCATTTAATCTTTAAGCTATTCAGAAATGAG 1620


```

QY 3781 AAGAACAAAGCCGAAATTCATCGAAGGCAATTAATGAGGATGCAAGTATC 3840
DB 3781 AAGAACAAAGCCGAAATTCATCGAAGGCAATTAATGAGGATGCAAGTATC 3840
QY 3841 AGTTCGCTTTAGGCAATTTGGCATTCAGCTTATTTGAGTTATCGCTATTTATG 3900
DB 3841 AGTTCGCTTTAGGCAATTTGGCATTCAGCTTATTTGAGTTATCGCTATTTATG 3900
QY 3901 AACGTGAAATTTATCATCTGAGAGAGTGAAGTGAAGAAACGCTGAGCTTGCATTAATC 3960
DB 3901 AACGTGAAATTTATCATCTGAGAGAGTGAAGTGAAGAAACGCTGAGCTTGCATTAATC 3960
QY 3961 GCTATATATGCTGCAATTCAGTTGATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
DB 3961 GCTATATATGCTGCAATTCAGTTGATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
QY 4021 AGCCTTATTTCTGCTGCAATTCAGTTGATTAATTAATTAATTAATTAATTAATTAAT 4080
DB 4021 AGCCTTATTTCTGCTGCAATTCAGTTGATTAATTAATTAATTAATTAATTAATTAAT 4080
QY 4081 ATCTACGAGTGTGCAACACCAATTCAGTTGATTAATTAATTAATTAATTAATTAAT 4140
DB 4081 ATCTACGAGTGTGCAACACCAATTCAGTTGATTAATTAATTAATTAATTAATTAAT 4140
QY 4141 CAGAAATTTTACATTTCCAAATTCCTTTTATCTGAAATTCGAAAGTTCAACTG 4200
DB 4141 CAGAAATTTTACATTTCCAAATTCCTTTTATCTGAAATTCGAAAGTTCAACTG 4200
QY 4201 GCAGAAATTTTACATTTCCAAATTCCTTTTATCTGAAATTCGAAAGTTCAACTG 4260
DB 4201 GCAGAAATTTTACATTTCCAAATTCCTTTTATCTGAAATTCGAAAGTTCAACTG 4260
QY 4261 ATCGTTATTTGATTAACAAAGTGGTCAATTCAGTTGATTAATTAATTAATTAATTAAT 4319
DB 4261 ATCGTTATTTGATTAACAAAGTGGTCAATTCAGTTGATTAATTAATTAATTAATTAAT 4319

```

RESULT 2

AD09342
ID AD09342 strand; DNA; 4319 BP.

AC
AD09342;

DT 06-NOV-2003 (first entry)

DE Haemophilus adherence and penetration protein (HAP) gene.

XX ds: Haemophilus adherence and penetration protein; HAP; gene;

KX antibacterial; immunostimulant; vaccine; Haemophilus influenza infection;

XX influenza.

OS Haemophilus influenzae.

XX US2003073166-A1.

ED 17-APR-2003.

PF 22-FEB-2002; 2002US-00080505.

XX 25-AUG-1994; 94US-00296791.

PR 20-APR-2001; 2001US-00839996.

XX (GENE/) GENE J W S.

XX GENE JWS;

XX WPI: 2003-567308/53.

XX P-ESDB; AD09343.

XX New recombinant Haemophilus adhesion and penetration protein, useful as a
vaccine for prophylactic or therapeutic treatment of a Haemophilus
influenzae infections, in generating antibodies for passive immunization.

PS Example 1; Fig 6; 149pp; English.
XX
CC The invention relates to a new recombinant Haemophilus adhesion and
CC penetration protein (HAP) appearing as AD09348, AD09350, AD09351, and
CC AD09354 encoded by a nucleic acid appearing as AD09349, AD09351, and
CC AD09353, AD09355 and AD09357. Also included are a recombinant nucleic
CC acid that will hybridize under stringency conditions to the nucleic acid
CC detailed above, an expression vector comprising transcriptional and
CC translation regulatory nucleic acid operably linked to the HAP nucleic
CC acid, producing HAP by culturing a host cell transformed with the
CC expression vector; a composition comprising a pharmaceutical carrier and
CC HAP, inducing an immune response in a patient by administering the
CC composition, a composition comprising an antibody capable of binding to
CC HAP (or to a peptide selected from 35 peptide sequences appearing as
CC AD09359-AD09393), and an antibody capable of binding to the peptides.
CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
CC infection (e.g. influenza), in generating an immune response in a
CC patient, and in generating antibodies to the HAP protein. The antibodies
CC may be used to diagnose the presence of H. influenzae infection in a
CC sample or patient, in passive immunisation treatments, and to purify or
CC separate HAP proteins or the H. influenzae organism from a sample. The
CC present sequence is the H. influenzae HAP gene.

SQ Sequence 4319 BP; 1497 A; 776 C; 891 G; 1155 T; 0 U; 0 Other;

Query Match 100.0%; Score 4319; DB 8; Length 4319;

Best Local Similarity 100.0%; Pred. No. 0; Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TCAATAGTGTGTTAATCTAGTATTTTAAATGCAAAATTAATTAATTAATTAATTAATTAAT 60
DB 1 TCAATAGTGTGTTAATCTAGTATTTTAAATGCAAAATTAATTAATTAATTAATTAATTAAT 60
QY 61 TGAATAAAAGCTGATTTGCTTAATTTTAAATGCAAAATTAATTAATTAATTAATTAATTAAT 120
DB 61 TGAATAAAAGCTGATTTGCTTAATTTTAAATGCAAAATTAATTAATTAATTAATTAATTAAT 120
QY 121 CGAGAGCGTGGCTGGTCACTATTTGGATTTGATTAATTAATTAATTAATTAATTAATTAAT 180
DB 121 CGAGAGCGTGGCTGGTCACTATTTGGATTTGATTAATTAATTAATTAATTAATTAATTAAT 180
QY 181 CGAGAGCGTGGCTGGTCACTATTTGGATTTGATTAATTAATTAATTAATTAATTAATTAAT 240
DB 181 CGAGAGCGTGGCTGGTCACTATTTGGATTTGATTAATTAATTAATTAATTAATTAATTAAT 240
QY 241 GGCATTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
DB 241 GGCATTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
QY 301 GTAAAGCGTGGCAGCCTTGTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
DB 301 GTAAAGCGTGGCAGCCTTGTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
QY 361 GATATACAGTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 420
DB 361 GATATACAGTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 420
QY 421 ATAAAGTTGTAAGAGAACTACATCAAAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 480
DB 421 ATAAAGTTGTAAGAGAACTACATCAAAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 481 ACCATTAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
DB 481 ACCATTAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
QY 541 ATATGAATGGGAGCTACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
DB 541 ATATGAATGGGAGCTACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
QY 601 CTGAGAGGAGTTTGGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
DB 601 CTGAGAGGAGTTTGGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660

```

QY 661 ATTATGACAGTGGCAATACACAAATAGGTGAGCAGGTAATGATATCTGATT 720
DB 661 ATTATGACAGTGGCAATACACAAATAGGTGAGCAGGTAATGATATCTGATT 720
QY 721 TGGAGGCGATGTTCTGTAAGCGGGAGAAATATGTCATTACCGATGCGCTCAAG 780
DB 721 TGGAGGCGATGTTCTGTAAGCGGGAGAAATATGTCATTACCGATGCGCTCAAG 780
QY 781 GGGACAGTGGTCTCGCATGTTATTTATGATGCTGAAAAAATAATGTTATATG 840
DB 781 GGGACAGTGGTCTCGCATGTTATTTATGATGCTGAAAAAATAATGTTATATG 840
QY 841 GGAATTAAGGGAGGCAACCTTTGAAAGCAAAATGAGTTTCAATGGTTGCA 900
DB 841 GGAATTAAGGGAGGCAACCTTTGAAAGCAAAATGAGTTTCAATGGTTGCA 900
QY 901 AATCTTATTTTATGTAATTTTCGAAAGATTTACATACATCACTTACACCGAGCTG 960
DB 901 AATCTTATTTTATGTAATTTTCGAAAGATTTACATACATCACTTACACCGAGCTG 960
QY 961 GTAATGGAATGTCACATTTAGTGAATGATATATGTCAGGGTCTATTAATCTCAAA 1020
DB 961 GTAATGGAATGTCACATTTAGTGAATGATATATGTCAGGGTCTATTAATCTCAAA 1020
QY 1021 CAGGAATACATCAGAAATTAATTAATACGTTAGCAAAATATGATGCTTGAAGAGA 1080
DB 1021 CAGGAATACATCAGAAATTAATTAATACGTTAGCAAAATATGATGCTTGAAGAGA 1080
QY 1081 AGATTAAGTTTATATCTATGATGATGAGGAACTTAATTTATTCCTCAGCTTTAAACA 1140
DB 1081 AGATTAAGTTTATATCTATGATGATGAGGAACTTAATTTATTCCTCAGCTTTAAACA 1140
QY 1141 ATGAGAAACGCTATATTTATGATCAAAAACAGGATCATTAATCTTGGATGCA 1200
DB 1141 ATGAGAAACGCTATATTTATGATCAAAAACAGGATCATTAATCTTGGATGCA 1200
QY 1201 TTAACCAAGGGCGGGTGGTCTTATTTGAGGTAATTTTACAGATCTCAAAATCTCA 1260
DB 1201 TTAACCAAGGGCGGGTGGTCTTATTTGAGGTAATTTTACAGATCTCAAAATCTCA 1260
QY 1261 ACCAACTTGGCAGAGAGCTGCAATCATGTAATGTAATGCAACCGTTACTGGAAAG 1320
DB 1261 ACCAACTTGGCAGAGAGCTGCAATCATGTAATGTAATGCAACCGTTACTGGAAAG 1320
QY 1321 TAAATGCGTGGAAATATGATGATTTTAAATTTGTAAGAAATATGCACTGTAAG 1380
DB 1321 TAAATGCGTGGAAATATGATGATTTTAAATTTGTAAGAAATATGCACTGTAAG 1380
QY 1381 CCAAAGGGGAAATTAAGTTTCAATGAGGTAGGCAATGTAATGCACTTGGAGAGC 1440
DB 1381 CCAAAGGGGAAATTAAGTTTCAATGAGGTAGGCAATGTAATGCACTTGGAGAGC 1440
QY 1441 AGGCAAGCATCAAGCAACAAACAGCTTTAGTAATGCTGTTAGCGCAGAG 1500
DB 1441 AGGCAAGCATCAAGCAACAAACAGCTTTAGTAATGCTGTTAGCGCAGAG 1500
QY 1501 GGAATCTTCAATTAACAGATTAACATTTGATACGATTAATTTATTTGCGCTTC 1560
DB 1501 GGAATCTTCAATTAACAGATTAACATTTGATACGATTAATTTATTTGCGCTTC 1560
QY 1561 GTGATGATGCTTAAATCTTAAAGGCAATTTCACTTAAAGGTAATGCAAAATACG 1620
DB 1561 GTGATGATGCTTAAATCTTAAAGGCAATTTCACTTAAAGGTAATGCAAAATACG 1620
QY 1621 ACAGAGGGGCAATGATTTGAAACCATTAACACTCAAGCGCTAAATGCTATTAAG 1680
DB 1621 ACAGAGGGGCAATGATTTGAAACCATTAACACTCAAGCGCTAAATGCTATTAAG 1680
QY 1681 GGAACGAAAGCATTTGTTACCTTAAGGAAATATTAATTAATTAATTAATTAAG 1740
DB 1681 GGAACGAAAGCATTTGTTACCTTAAGGAAATATTAATTAATTAATTAATTAAG 1740

QY 1741 AAATGCTCAACAGGTTGGTTGGCAAAACAGATTAATAATAACAAATGGGCAATTA 1800
DB 1741 AAATGCTCAACAGGTTGGTTGGCAAAACAGATTAATAATAACAAATGGGCAATTA 1800
QY 1801 ACCTTATTTATTAACCAACCAAGAAATGCTTCTGCTTCACTTCAAGTGTGTAATTA 1860
DB 1801 ACCTTATTTATTAACCAACCAAGAAATGCTTCTGCTTCACTTCAAGTGTGTAATTA 1860
QY 1861 TAAAGGCGATTAATCCCAAAAGGTAACATTTTTCAGGTTAGCCGACCGC 1920
DB 1861 TAAAGGCGATTAATCCCAAAAGGTAACATTTTTCAGGTTAGCCGACCGC 1920
QY 1921 AGGCTACATCATTTTAAATTAACGTTGTCAGAAATGAGATTAACCAAGGCGAA 1980
DB 1921 AGGCTACATCATTTTAAATTAACGTTGTCAGAAATGAGATTAACCAAGGCGAA 1980
QY 1981 TTGTGGGATCAGATTTGATCAACGTTACATTTAAAGTGAATCTTCAATTAAG 2040
DB 1981 TTGTGGGATCAGATTTGATCAACGTTACATTTAAAGTGAATCTTCAATTAAG 2040
QY 2041 GGGAGAGTGGGTTGTTCTGCAATGTTTCTCAATGAGGAAATTTGACATGACGA 2100
DB 2041 GGGAGAGTGGGTTGTTCTGCAATGTTTCTCAATGAGGAAATTTGACATGACGA 2100
QY 2101 ATATGCAAAATGCAATTTGTTGTTGCAATCAACAAATCCATTTGACGGGT 2160
DB 2101 ATATGCAAAATGCAATTTGTTGTTGCAATCAACAAATCCATTTGACGGGT 2160
QY 2161 CAGATTGACAGATTAAGCACTTGTCAAAAATGATTTAAACGATTAACAAATGATTA 2220
DB 2161 CAGATTGACAGATTAAGCACTTGTGTCAAAAATGATTTAAACGATTAACAAATGATTA 2220
QY 2221 ATTCTATACCAAAACCAATCAATGCTGCTATTTTAACTGATTAAGCAACGGCA 2280
DB 2221 ATTCTATACCAAAACCAATCAATGCTGCTATTTTAACTGATTAAGCAACGGCA 2280
QY 2281 ATGTTAAAGTTTGAACAACTTAATGCAATGTCATTAAACAAATCAGCCAAATTA 2340
DB 2281 ATGTTAAAGTTTGAACAACTTAATGCAATGTCATTAAACAAATCAGCCAAATTA 2340
QY 2341 CATTAAGCAACATGCAACCAATGCAATTTGCACTTTCGCAATTTCAATGCA 2400
DB 2341 CATTAAGCAACATGCAACCAATGCAATTTGCACTTTCGCAATTTCAATGCA 2400
QY 2401 CGGTGATTAAGCAATGCAATGCAATTTGCACTTTCGCAATTTCAATGCA 2460
DB 2401 CGGTGATTAAGCAATGCAATGCAATTTGCACTTTCGCAATTTCAATGCA 2460
QY 2461 TAAAAAACGCAATTTTGCACCAATTTGAGGAGCAAAAGCAACAGTACGTTG 2520
DB 2461 TAAAAAACGCAATTTTGCACCAATTTGAGGAGCAAAAGCAACAGTACGTTG 2520
QY 2521 AAATGCACTTGGCAATGCTAGCATTAATGCAATTTGCAATTTAAACGTTAAATCA 2580
DB 2521 AAATGCACTTGGCAATGCTAGCATTAATGCAATTTGCAATTTAAACGTTAAATCA 2580
QY 2581 GTCATGATGATTAATGCAATTTGCACTTTCAGCTGCTCAAAATGCAAGTGGCGTT 2640
DB 2581 GTCATGATGATTAATGCAATTTGCACTTTCAGCTGCTCAAAATGCAAGTGGCGTT 2640
QY 2641 CATTAAGAGCGAAACCAACGCAATGCAATGCAATGCTTCAACATTTGACATTA 2700
DB 2641 CATTAAGAGCGAAACCAACGCAATGCAATGCAATGCTTCAACATTTGACATTA 2700
QY 2701 ATGATTAATTAATGAGGCAAGGCAATTTCAATTTTACTTATTTGCTATTAATA 2760
DB 2701 ATGATTAATTAATGAGGCAAGGCAATTTCAATTTTACTTATTTGCTATTAATA 2760
QY 2761 GGAATTAATTAATTTTCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAAT 2820
DB 2761 GGAATTAATTAATTTTCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAAT 2820
QY 2821 CAGGCAAGAAACCGCAACCTTTGAGCAATTAATTTGTTGAAAGCAAAATTAATCAAC 2880

```

Db      2821  CAGGCAAGACCCCAAAACCTTGACAAATTAATTTGGTGAAGCAAGATTAATCAAC 2880
Qy      2881  CGTATCAGTAAGTCAATTAATTTGAGAAAATGACACAGTGTATGAGGTGATTAAC 2940
Db      2881  CGTATCAGTAAGTCAATTAATTTGAGAAAATGACACAGTGTATGAGGTGATTAAC 2940
Qy      2941  GTTATTAATTAAGTAAGATGATGAGCAATTCGGCTTGCATTAACCCATTAAGAGCAG 3000
Db      2941  GTTATTAATTAAGTAAGATGATGAGCAATTCGGCTTGCATTAACCCATTAAGAGCAG 3000
Qy      3001  AATTGCAATATGATTTAGTAAGACAGACAGCAAGCAAGCAATTAAGAGCAGCAAG 3060
Db      3001  AATTGCAATATGATTTAGTAAGACAGACAGCAAGCAAGCAATTAAGAGCAGCAAG 3060
Qy      3061  TTGACCGACTGTGTAACCAACAGAGTGAAGCAAGAGTGGTGAAGAGAGAGAGCA 3120
Db      3061  TTGACCGACTGTGTAACCAACAGAGTGAAGCAAGAGTGGTGAAGAGAGAGAGCA 3120
Qy      3121  GAGCAGCGTTCCTGATACCTGCTGATCAAGAGCTGTTAAAGCATTGAGAGCAAC 3180
Db      3121  GAGCAGCGTTCCTGATACCTGCTGATCAAGAGCTGTTAAAGCATTGAGAGCAAC 3180
Qy      3181  AAGTGAACCTGACCTGTAACCAACCAAGTAAGCAAGCAAGCAAGCAAGCAAGCA 3240
Db      3181  AAGTGAACCTGACCTGTAACCAACCAAGTAAGCAAGCAAGCAAGCAAGCAAGCA 3240
Qy      3241  AAGAGCAGTGTTCCTGATACCTGCTGATCAAGAGCTGTTAAAGCATTGAGAGCA 3300
Db      3241  AAGAGCAGTGTTCCTGATACCTGCTGATCAAGAGCTGTTAAAGCATTGAGAGCA 3300
Qy      3301  TTGAGTTATTAATGACCCCAACAGCAATGCAAGCAAGCAAGCAAGCAAGCAAGCA 3360
Db      3301  TTGAGTTATTAATGACCCCAACAGCAATGCAAGCAAGCAAGCAAGCAAGCAAGCA 3360
Qy      3361  AAAAAACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3420
Db      3361  AAAAAACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3420
Qy      3421  TATCTGCAACAGTAATGATGATGCTTCTGCTGATCAAGATGATGATGATGATGAT 3480
Db      3421  TATCTGCAACAGTAATGATGATGCTTCTGCTGATCAAGATGATGATGATGATGAT 3480
Qy      3481  ATCAAGCAACATTCGCTGATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3540
Db      3481  ATCAAGCAACATTCGCTGATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3540
Qy      3541  ATGCGTTCGCTGATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3600
Db      3541  ATGCGTTCGCTGATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3600
Qy      3601  CTTAGCTAATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3660
Db      3601  CTTAGCTAATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3660
Qy      3661  ATGAACAGGTTAAATTAATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3720
Db      3661  ATGAACAGGTTAAATTAATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3720
Qy      3721  GGGGCGATTGACATTTGCTTAAGCTGGAAGCGGAAATGAGCGAGTAATATGCTG 3780
Db      3721  GGGGCGATTGACATTTGCTTAAGCTGGAAGCGGAAATGAGCGAGTAATATGCTG 3780
Qy      3781  AAGAACAAGCCGAAATTAATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3840
Db      3781  AAGAACAAGCCGAAATTAATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3840
Qy      3841  AGTTCGCTTTGAGCAATTTGGGCAATTCAGCTTAATTTGAGATTAATCGCTATTT 3900
Db      3841  AGTTCGCTTTGAGCAATTTGGGCAATTCAGCTTAATTTGAGATTAATCGCTATTT 3900
Qy      3901  AACGTGAATTAATCAATCTGAGAGAGTGAAGTGAAGCAAGCAAGCAAGCAAGCA 3960

```

```

Db      3901  AACGTGAATTAATCAATCTGAGAGAGTGAAGTGAAGCAAGCAAGCAAGCAAGCA 3960
Qy      3961  GCATTAATGCTGATGATGATGATTAATCAATTAATCAATTAATCAATTAATCA 4020
Db      3961  GCATTAATGCTGATGATGATGATTAATCAATTAATCAATTAATCAATTAATCA 4020
Qy      4021  AGCTTAATTTCTGCTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 4080
Db      4021  AGCTTAATTTCTGCTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 4080
Qy      4081  ATCTCAAGCTGCTGACCAACCAATTTGACAGCTTAATTTGAGCAATTTGAGCA 4140
Db      4081  ATCTCAAGCTGCTGACCAACCAATTTGACAGCTTAATTTGAGCAATTTGAGCA 4140
Qy      4141  CAGAAATTTTCAATTTCAATTTCCGCTTATTAATTTCAATTTCAATTTCAATTT 4200
Db      4141  CAGAAATTTTCAATTTCAATTTCCGCTTATTAATTTCAATTTCAATTTCAATTT 4200
Qy      4201  GCAACAGCAAAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4260
Db      4201  GCAACAGCAAAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4260
Qy      4261  ATGCTTAATTAATTAACAGAGTGGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4319
Db      4261  ATGCTTAATTAATTAACAGAGTGGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4319

```

RESULT 3

ADA09349
ID ADA09349 standard; DNA; 4305 BP.

ADA09349;

06-NOV-2003 (first entry)

Haemophilus HMP protein gene from strain 11.

ds; Haemophilus adherence and penetration protein; HAP; gene;

antibacterial; immunostimulant; vaccine; Haemophilus influenza infection;

Haemophilus influenzae; strain 11.

US2003073166-A1.

17-APR-2003.

22-FEB-2002; 2002US-03080505.

25-AUG-1994; 94US-00296791.

20-APR-2001; 2001US-00839996.

(GEMV/) GEMV J W S.

Geme JWS;

WPI; 2003-567308/53.

P-PsDB; ADA09350.

New recombinant Haemophilus adhesion and penetration protein, useful as a

vaccine for prophylactic or therapeutic treatment of a Haemophilus

influenzae infections, in generating antibodies for passive immunization.

Claim 1; Fig 16; 143pp; English.

The invention relates to a new recombinant Haemophilus adhesion and
penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
acid that will hybridize under stringency conditions to the nucleic acid
described above, an expression vector comprising transcriptional and
translation regulatory nucleic acid operably linked to the HAP nucleic
acid, producing HAP, by culturing a host cell transformed with the

CC expression vector, a composition comprising a pharmaceutical carrier and
 CC HAP, inducing an immune response in a patient by administering the
 CC composition, a composition comprising an antibody capable of binding to
 CC HAP (or to a peptide selected from 35 peptide sequences appearing as
 CC ADA09359-ADA09393), and an antibody capable of binding to the peptides.
 CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
 CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
 CC infection (e.g. influenza), in generating an immune response in a
 CC patient, and in generating antibodies to the HAP protein. The antibodies
 CC may be used to diagnose the presence of H. influenzae infection in a
 CC sample or patient, in passive immunisation treatments, and to purify or
 CC separate HAP proteins or the H. influenzae organism from a sample. The
 CC present sequence is a HAP gene from a strain of H. influenzae.
 XX
 SQ Sequence 4305 BP; 1490 A; 762 C; 902 G; 1150 T; 0 U; 1 Other;
 Query Match 73.6%; Score 3180.8; DB 8; Length 4305;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 3704; Conservative 0; Mismatches 463; Indels 156; Gaps 9;

QY 60 ATGAAAAAAGTGTATTTGCTTAAATTTTAAACGGCTTGCAATTCATTAGAGATGTA 119
 DB 1 ATGAAAAAAGTGTATTTGCTTAAATTTTAAACGGCTTGCAATTCATTAGAGATGTA 60
 QY 120 TCGAAGCGTGGCTGTCACATTTTGGATTGATTACAAATTAATTCGTTATTT 179
 DB 61 TCGAAGCGTGGCTGTCACATTTTGGATTGATTACAAATTAATTCGTTATTT 120
 QY 180 GCGGAAATTAAGAGAGATTCACAGTTGGGGCTCAAAATATTAAGTTTATACAAACA 239
 DB 121 GCGGAAATTAAGAGAGATTCACAGTTGGGGCTCAAAATATTAAGTTTATACAAACA 180
 QY 240 GGGCAATTAAGTGGCAATCATGACAAAGCCCGATGATTTTCTAGATGCA 299
 DB 182 GGGCAATTAAGTGGCAATCATGACAAAGCCCGATGATTTTCTAGATGCA 240
 QY 300 CGTAAAGCGCTGGCAAGCTTGGTGAATCAATTAATTTGAGCGTGGCAATACGTA 359
 DB 241 AGAATGAGAGTTCCTGCTTAGTAGGCGATCAGTATATTGAGTGGCAATACGTA 300
 QY 360 GGAATACAGATTCGATTTTGTGCGAGGAGAAACAACCCGATCAACATCGTTTACT 419
 DB 301 GGAATACAGATTCGATTTTGTGCGAGGAGAAACAACCCGATCAACATCGTTTACT 360
 QY 420 TATAGATTGTAAACGAAATTAACCAAAAAGATTAATTAACCTTATAGAGACAT 479
 DB 361 TATAGATTGTAAACGAAATTAACCAAAAAGATTAATTAACCTTATAGAGACAT 420
 QY 480 TACCAATATCAAGATTAACATTAATTTGTTAGAGAGCGCTCAATGATATGAC 539
 DB 421 TACCAATATCAAGATTAACATTAATTTGTTAGAGAGCGCTCAATGATATGAC 480
 QY 540 AATATGAATGGAGATTCATTAAGATTAAGAAACAAATTCAGAAAGTGTGATGCGC 599
 DB 481 AATATGAATGGAGATTCATTAAGATTAAGAAACAAATTCAGAAAGTGTGATGCGC 540
 QY 600 TCTGAGCGAGTTTGGCGAATGATCAAGCAAGGCGACCAAGTTCCCGGTGCATAT 659
 DB 541 TCTGAGCGAGTTTGGCGAATGATCAAGCAAGGCGACCAAGTTCCCGGTGCATAT 600
 QY 660 CATTATCGACAGCTGGCAATTAACACATTCAGCGTGGAGAGAGATGATATTCGAT 719
 DB 601 AATATTTTAAAGTGGAGATTAACATTAATTAAGGAGAGATGATATTCGAT 660
 QY 720 TTGGAGAGCGATGTTCTGAAGCGGAGAAATGTCATTAACGATTCGAGGTCGAAG 779
 DB 661 GTGAGAGCGATGTTCTGAAGCGGAGAAATGTCATTAACGATTCGAGGTCGAAG 720
 QY 780 GGGGACAGTGGTCTCCGATGTTTATTTATGATGCTGAAGCAAAATATGTTATTAAT 839
 DB 721 GGGGACAGTGGTCTCCGATGTTTATTTATGATGCTGAAGCAAAATATGTTATTAAT 780
 QY 840 GGGATTTATGAGGAGCAAGCCCTTTTGAAGGCAAGAAATGGGTTTCAATTTGCTGC 899

DB 781 GGGATTTGCGGGGGGCAACCTTATACAGGCAATTCAGATGCTTCAATTAAGCCCT 840
 QY 900 AATCTTATTTGATGAAATTTTTCAGAGATTTGATATATCATCTTACACCGAGCT 959
 DB 841 AATCTTATTTGATGAAATTTTTCAGAGATTTGATATATCATCTTACACCGAGCT 900
 QY 960 GGTATGAGATGATACCAATTAATGAGAAATGATTAAGTTCAGAGGCTATTAACAGAAA 1019
 DB 901 GGTATGAGATGATTAACATTTTCAATTTCAAGATGATTAAGTTCAGAGGCTATTAAC 960
 QY 1020 TCGAATTCATCAAGAAATTAATTAATTAAGTTCAGAAATTAAGTTCAGAAAGAG 1079
 DB 961 ATTGGAAGAAC-----TCGTGAATTAACCTTTACGAAACAAATTAAGAAAT 1014
 QY 1080 AAGGATTAAGTCAATTAATCTGATTAAGAGAGCACTTAATTTATCTCCAGCTTAAC 1139
 DB 1015 AAGGATTAAGTCAATTAATCTGATTAAGAGAGCACTTAATTTATCTCCAGCTTAAC 1074
 QY 1140 AATGAGAAAGCGTATATTTATGATCAAAACAGAGATCATTAATCTGATGAC 1199
 DB 1075 AATGAGAAAGATTTATTTTGGAGATGAAGATTAAGATCATTAATCTGATGAC 1134
 QY 1200 ATTACCAAGAGCGGCTGCTTATTTTGAAGGATTAATTTACGATCTCAATTC 1259
 DB 1135 ATGATCAAGGATGAGCGGCTTGTATTTTGAAGGAGATTTATAGTTCCCTACCAAA 1194
 QY 1260 AACCAACCTTGGCAAGAGCTGCGATACATGATTAAGAAATTAAGACGCTACTGAAA 1319
 DB 1195 AATGAAGGTGAAGAGTGGGCAATCAATGATGAGAAATTAAGACGCTACTGAAA 1254
 QY 1320 GTAAATGCGTGAACATGATGACATCTTCTAAATTTGTTAAGGAACTTGACCTCA 1379
 DB 1255 GTAAATGCGTGAACATGATGACATCTTCTAAATTTGTTAAGGAACTTGACCTCA 1314
 QY 1380 GCCAAAGGGAAATTAAGGTTGATTCGCTGATGAGGATGATTAAGGATTTTGAAGCAG 1439
 DB 1315 GCCAAAGGGAAATTAAGGTTGATTCGCTGATGAGGATGATTAAGGATTTTGAAGCAG 1374
 QY 1440 CAGGAGAGATCAAGGCAACAAACAGCTTTAGGAAATTTGGCTGATGAGCGAG 1499
 DB 1375 CAGGAGAGATCAAGGCAACAAACAGCTTTAGGAAATTTGGCTGATGAGCGAG 1434
 QY 1500 GGGACGTTCAATTAACGATTAACCAATTTGATTAACGATTAATTTTATTTGAGCTT 1559
 DB 1435 GGGACGTTCAATTAACGATTAACCAATTTGATTAACGATTAATTTTATTTGAGCTT 1494
 QY 1560 CGTGGTGGTGGTATGATCTTAAGCGGATTCATTAACCTTTAAGGATTCGAATACG 1619
 DB 1495 CGTGGTGGTGGTATGATCTTAAGCGGATTCATTAACCTTTAAGGATTCGAATACG 1554
 QY 1620 GAGGAGGGGCAATGATTTGAAACCATTAATCAATCAAGCGCTAATGCTACTTACT 1679
 DB 1555 GAGGAGGGGCAATGATTTGAAACCATTAATCAATCAAGCGCTAATGCTACTTACT 1614
 QY 1680 GGGAGCAAGACAT---TGTTCTACTATGAGAAATTAATTAATTAACCTTGAATCA 1736
 DB 1615 GGGAGCAAGACATTTACTGCTCATCTAATAAATAATTAATTAACCTTGAATCA 1674
 QY 1737 AAGGAATTTGCCCAAGGTTGTTTGGGCAACAGATTAATTAATTAACCAATTAAGGCGA 1796
 DB 1675 AAGGAATTTGCCCAAGGTTGTTTGGGCAACAGATTAATTAATTAACCAATTAAGGCGA 1734
 QY 1797 TTAACCTTATTTAATTAACCAACAGAGATGATGATGCTTCACTTCAAGTGTGATCA 1856
 DB 1735 TTAACCTTATTTAATTAACCAACAGAGATGATGATGCTTCACTTCAAGTGTGATCA 1794
 QY 1857 AATTTAAAGCGATTAATTAACCAACAGAGATTAATTTTGAAGGATTAATTAACCAAGC 1916
 DB 1795 AATTTAAAGCGATTAATTAACCAACAGAGATTAATTTTGAAGGATTAATTAACCAAGC 1854
 QY 1917 CCGGAGGCTTACATCTTAATTAATTAAGTGTGCGAAATTAAGGATTAATTAACCAAGC 1976

1855 CCCCACCCCTACATCATTTAGACAAAGCTTGGTCAGAAATGAGAGGTATCCCAACAGGC 1914
QY 1977 GAATTTGTGGATACGATTTGATGACACCGATCACTTTAAGTGAAGAACTCCAAAT 2036
Db 1915 GAAATTTGTGGATTAAGATTTGATTTAGCCGATTTAAGCTGAAACTTCCAAAT 1974
QY 2037 AAAAGCGGAAGTGGCTGGTTTCTCGCATGTTCTCAATTGAGGAAATGGAAGTC 2096
Db 1975 AAAAGCGGAAGTGGCTGGTTTCTCGCATGTTCTCAATTGAGGAAATGGAAGTC 2034
QY 2097 AGCAATATGCAAAATGCAATTTGGTGTGGCAAAATCAAAATTCATTTGACG 2156
Db 2035 AGCAATATGCAAAATGCAATTTGGTGTGGCAAAATTCATTTGACG 2094
QY 2157 CGTTCAGATTGACAGATTAAGCATGTCGCAAAAGTGAATTAAACGATCAAAAGTT 2216
Db 2095 CGTTCAGATTGACAGATTAAGCATGTCGCAAAAGTGAATTAAACGATCAAAAGTT 2154
QY 2217 ATTAATTTATACCAAAACACAAATCATGCTCTATTAATTAATGATATGACAG 2276
Db 2155 ATTAATTTATACCAAAACACAAATCATGCTCTATTAATTAATGATATGACAG 2214
QY 2277 GCGAATTTAAAGTTTACCAAACTTAATGCAATGTCATTTAAACAAATGACGCA 2336
Db 2215 GCGAATTTAAAGTTTACCAAACTTAATGCAATGTCATTTAAATATATGACGCA 2274
QY 2337 TTTCATTTAGCAACATGCAACCAATATGCAATGCAATTTGACGCAATTCAGCT 2396
Db 2275 TTTCATTTAGCAACATGCAACCAATATGCAATGCAATTTGACGCAATTCAGCT 2234
QY 2397 ----- 2396
Db 2335 GCAAGGTAATATGATGCACTTTAATGGGATGATTTAGGCAATGATGCGGTTT 2294
QY 2397 -----GCA 2399
Db 2395 ACATTAAGCAATCAAGCAACAGATGCAATGCAATGCAATGCAATGCAATGCA 2454
QY 2400 ACATTAAGCAATCAAGCAACAGATGCAATGCAATGCAATGCAATGCAATGCA 2459
Db 2455 ACATTAAGCAATCAAGCAACAGATGCAATGCAATGCAATGCAATGCAATGCA 2414
QY 2460 TTAATAAAGCAATTTTGGCAAAATTCAGGAGCAAAAGCAACAGATGCAATGCA 2519
Db 2515 TTAATAAAGCAATTTTGGCAAAATTCAGGAGCAAAAGCAACAGATGCAATGCA 2574
QY 2520 GAAATGCGATTTGCAAAATTCAGGAGCAAAAGCAACAGATGCAATGCAATGCA 2579
Db 2575 GAAATGCGATTTGCAAAATTCAGGAGCAAAAGCAACAGATGCAATGCAATGCA 2534
QY 2580 AGTACGATCAGTTAAATTTGATTTAGCTCAAAATGCAATGCAATGCAATGCA 2636
Db 2635 AGTACGATCAGTTAAATTTGATTTAGCTCAAAATGCAATGCAATGCAATGCA 2634
QY 2637 CGTTCATTTAGCAACAGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2696
Db 2695 CGTTCATTTAGCAACAGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2754
QY 2697 GATTAATGCAATTTAGGAGGCAAGCAATTTGCAATTTGCAATTTGCAATTTGCA 2756
Db 2755 GATTAATGCAATTTAGGAGGCAAGCAATTTGCAATTTGCAATTTGCAATTTGCA 2814
QY 2757 AAAAGCATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2816
Db 2815 AAAAGCATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2874
QY 2817 AAACAGGCAAAAGCAACAGATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2876
Db 2875 AAACAGGCAAAAGCAACAGATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2934
QY 2877 CAACGTTATGCAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2936
Db 2935 CAACGTTATGCAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2994

QY 2937 TTACGTTAATTAATTTAGGAGAAATGCAATGCAATTTGCAATTTGCAATTTGCA 2996
Db 2995 TTACGTTAATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3054
QY 2997 CAGAAATTTGCAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3056
Db 3055 CAGAAATTTGCAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3114
QY 3057 CAGAAATTTGCAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3116
Db 3115 CAGAAATTTGCAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3168
QY 3117 GCGAGAGCAAGGTTTCTGATACCTGCTGATCAAGCTGATCAAGCTGATCAAGCT 3176
Db 3169 ---AGAGGAGGTTTCTGATACCTGCTGATCAAGCTGATCAAGCTGATCAAGCT 3225
QY 3177 AAACAGGCAAAAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3236
Db 3226 AAACAGGCAAAAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3285
QY 3237 TCAAAAAGGCT-----AGTGTTCGATCCCTGCTGATCAAGCTGATCAAGCT 3284
Db 3286 TCAAAAAGGCT-----AGTGTTCGATCCCTGCTGATCAAGCTGATCAAGCT 3345
QY 3285 GCAATTAAGGAGGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3344
Db 3346 GCAATTAAGGAGGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3405
QY 3345 GCTCAA---GAAAGGCAAAAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3401
Db 3406 GCTCAA---GAAAGGCAAAAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3465
QY 3402 AAATGCGATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3461
Db 3466 AAATGCGATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3525
QY 3462 TTACGTTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3521
Db 3526 TTACGTTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3585
QY 3522 AAAGAGGCTTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3581
Db 3586 AAAGAGGCTTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3642
QY 3582 CAAATTTGAGGAGGCAAAAGCTTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCA 3641
Db 3643 CAAATTTGAGGAGGCAAAAGCTTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCA 3702
QY 3642 CGTTCAGATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3701
Db 3703 CGTTCAGATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3762
QY 3702 GGTTCGCAATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3761
Db 3763 GGTTCGCAATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3822
QY 3762 AGTGCAGATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 3821
Db 3823 AGTGCAGATTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3882
QY 3822 GAGGTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3881
Db 3883 GAGGTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3942
QY 3882 GTTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3941
Db 3943 GTTAATTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 4002
QY 3942 CTAAGGCTTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 4001
Db 4003 CTAAGGCTTTAGGAGAAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 4062

1332 GGAACATGATCGACATTCATAAAATGGTAAGGACATTCACCGTCCAGCCAAAGGGGGA1333

QY 1392 AATTAAGGTTGATGACGGGTAGGAGTGTAAAGTCAATTTTGAGCGACGACGAT 1451
DB 77147 AATTAGGAAGTTAAGTGGGGTACGCGAAGGATCTTATGATCAACAGAGATGAA 77088
QY 1452 CAAGCGAACAACACCTTTAGTAAATTTGGCTGTGTACGCGCAAGGACCTTTCA 1511
DB 77087 GCGGGTCAAAAACACCTTTCAAAAGTTGGATGTGAACGGTTCAGAACATGTCA 77028
QY 1512 TTAAACGATGATTAACATTTGATACCGATTAATTTTGTGGCTTGGTGGTGGC 1571
DB 77027 TTAAATAGTACAGTGAATGATGATCTTAACAATCTATTTGGATTTGGTGGTGGC 76968
QY 1572 TTAGTCTTAACGCGGATTCATTAACCTTTTAAAGTATCCAAATTAACGACGAGGCG 1631
DB 76967 TTAGTCTTAACGCGGATTCATTAACCTTTTAAAGTATCCAAATTAACGAGGCGG 76908
QY 1632 ATGATGTGAACATTAATCACTCAAGCGGCTAATGTCATCTATTTACGGAACCAAGC 1691
DB 76907 ATGATGTGAACATTAATCACTCAAGTGTAAATATTAATTAATTAATTAATTA 76848
QY 1692 AT--TGTCTACCTAATGSAATTAATTAATTAATTAATTAATTAATTAATTA 1748
DB 76847 ATTAATGCTCATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 76788
QY 1749 TACAACGCTTGTGTGGGAAACGATTAATTAATTAATTAATTAATTAATTAATTA 1808
DB 76787 TACAACGCTTGTGTGGGAAACGATTAATTAATTAATTAATTAATTAATTAATTA 76728
QY 1809 TATAACCAACACAGAGATGATCTTACTTACTTACTTACTTACTTACTTACTTACT 1868
DB 76727 TATAACCAACACAGAGATGATCTTACTTACTTACTTACTTACTTACTTACTTACT 76668
QY 1869 GATAATTAACCAACAGAGATGATCTTACTTACTTACTTACTTACTTACTTACTTACT 1928
DB 76667 GATAATTAACCAACAGAGATGATCTTACTTACTTACTTACTTACTTACTTACTTACT 76638
QY 1929 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1988
DB 76607 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 76548
QY 1989 GATCAAGTATGATCAACGCTTACTTACTTACTTACTTACTTACTTACTTACTTACT 2048
DB 76547 GATCAAGTATGATCAACGCTTACTTACTTACTTACTTACTTACTTACTTACTTACT 76488
QY 2049 GCGGTGTCTTCTGCAATGTTTCTTCAATTAAGGAAATTTGACAGTCAATTAATCA 2108
DB 76487 GCGGTGTCTTCTGCAATGTTTCTTCAATTAAGGAAATTTGACAGTCAATTAATCA 76428
QY 2109 AATGCAATTTGT 2168
DB 76427 AATGCAATTTGT 76368
QY 2169 ACAGGATTAACGATCTGTCAAAAGTGTGAATTTTAAACGATTAATTAATTAATTA 2228
DB 76367 ACAGGATTAACGATCTGTCAAAAGTGTGAATTTTAAACGATTAATTAATTAATTA 76308
QY 2229 CCAAAAACAAAATCAATGCTCTATTAATTAATTAATTAATTAATTAATTAATTAAT 2288
DB 76307 CCAAAAACAAAATCAATGCTCTATTAATTAATTAATTAATTAATTAATTAATTAAT 76248
QY 2289 GGTTTAGCAAACTTAATGCAATGTCATTTAACAATCAACGACCAATTTCAATTAAGC 2348
DB 76247 GGTTTAGCAAACTTAATGCAATGTCATTTAACAATCAACGACCAATTTCAATTAAGC 76188
QY 2349 AACATGCAACCAATTAAGCAATATTCGATTTCCGACATTTCAATGCAACGCTGAT 2408
DB 76187 AACATGCAACCAATTAAGCAATATTCGATTTCCGACATTTCAATGCAACGCTGAT 76128
QY 2409 AATGCAACCTTGAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
DB 76127 AATGCAACCTTGAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 76068
QY 2469 AGGCAATTTTGCACCAAAATTCAGGAGACAAAGGCAACAGTGAAGCTTGAAGAAATGCG 2528

DB 76067 AGGCAATTTTGCACCAAAATTCAGGAGCAAAAGCAACAGTGAAGCTTGAAGAAATGCG 76008
QY 2529 ACTTGCAAAAGCTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2588
DB 76007 ACTTGCAAAAGCTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 75948
QY 2589 ACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2645
DB 75947 ACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75888
QY 2646 GAGACGGAACCAAGCCATTCGACGACGATGATGATGATGATGATGATGATGATGATGAT 2705
DB 75887 GAGACGGAACCAAGCCATTCGACGACGATGATGATGATGATGATGATGATGATGAT 75828
QY 2705 AATTAAGTGGGCAAGGCAATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2765
DB 75827 AATTAAGTGGGCAAGGCAATTCGATTAATTAATTAATTAATTAATTAATTAATTAAT 75768
QY 2766 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2825
DB 75767 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 75708
QY 2826 AAGAACCCGGAACCTTGAACGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2885
DB 75707 AAGAACCCGGAACCTTGAACGATTAATTAATTAATTAATTAATTAATTAATTAATTA 75648
QY 2886 TCAGATTAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2945
DB 75647 TCAGATTAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 75588
QY 2946 AATTAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3005
DB 75587 AATTAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75528
QY 3006 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3065
DB 75527 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 75468
QY 3066 CCACTGTCTAAACCAAAACAGGTGACCAAAAGTGGGTCAAGAAAGGACGACGACGAC 3125
DB 75467 CCACTGTCTAAACCAAAACAGGTGACCAAAAGTGGGTCAAGAAAGGACGACGACGAC 75407
QY 3126 GCGTTTCTGATACCTGCTGATCAAAAGCTGTAAAGCAATTAAGCAACGATGAT 3185
DB 75407 GCGTTTCTGATACCTGCTGATCAAAAGCTGTAAAGCAATTAAGCAACGATGATGAT 75347
QY 3186 GAATGATCTGTGAACCAAAAGTGAAGCAAAACCAAAAGTGGGTCAAAAGCA 3245
DB 75346 GAATGATCTGTGAACCAAAAGTGAAGCAAAACCAAAAGTGGGTCAAAAGCA 75297
QY 3246 -----GAGTGTCTTGTGATCCCTGCTGATCAAAAGCTGT-----GCAATTAAGAA 3293
DB 75296 GCAATGATGACGCTTGTGATCCCTGCTGATCAAAAGCTGT-----GCAATTAAGAA 75237
QY 3293 GCGGCACTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3353
DB 75236 GCGGCACTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75177
QY 3354 GAAGCGGAACCAAGCAACCAAAAGTGAAGCAAAACCAAAAGTGGGTCAAAAGCA 3413
DB 75176 G-----AACAGGCAACCAAAAGTGAAGCAAAACCAAAAGTGGGTCAAAAGCA 75126
QY 3414 TCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3473
DB 75125 TCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 75066
QY 3474 TTTGTGATGAAGCAATTTGCTGCTGTGTGAACAAATTCGACAGGATTAAGGACCTTAT 3533
DB 75065 TTTGTGATGAAGCAATTTGCTGCTGTGTGAACAAATTCGACAGGATTAAGGACCTTAT 75006
QY 3534 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3593

QY 2556 ACTGACATTCGAGATTTTAAGCTTAATAAGATGACGATGCTTAATTCAGCTTTATCA 2609
 Db 2944 GCGACATTCGAGATTTTAAGCTTAATAAGATGACGATGCTTAATTCAGCTTTATCA 3003
 QY 2610 GCTAGCTCAAAATATGAGCCAGCT---CGCCGTTCAATTAAGACGAAACCAACCCACCA 2666
 Db 3004 GCTAGCTCAAAATATGAGCCAGCTCAACCCGTTCAATTAAGACGAAACCAACCCACCA 3063
 QY 2667 TCGGACGAACATCGTTTCAACATTTGACGTAATGTAATGAGTGCGGCAAGGCA 2726
 Db 3064 TCGGACGAACATCGTTTCAACATTTGACGTAATGTAATGAGTGCGGCAAGGCA 3123
 QY 2727 TTCCATTTACTCCATCTTAATTTGGCTTAATAAGCATTAATTAATTTCCATGAC 2786
 Db 3124 TTCCATTTACTCCATCTTAATTTGGCTTAATAAGCATTAATTAATTTCCATGAC 3183
 QY 2787 GCTGAGGCGCATTAATATATCTGTTGCAACACAGGCAAAAGACCCGAACCCCTTGAG 2846
 Db 3184 GCTGAGGCGCATTAATATATCTGTTGCAACACAGGCAAAAGACCCGAACCCCTTGAG 3243
 QY 2847 CAATTAATCTTGTGAAAGCAAAATATCAACCCGTTATCAGATTAAGCTCAATTTACT 2906
 Db 3244 CAATTAATCTTGTGAAAGCAAAATATCAACCCGTTATCAGATTAAGCTCAATTTACT 3303
 QY 2907 TTGAAATGACCAAGCTTAATGAGGTGCACTTAAGCTTAATTAATTAATTAATTAAT 2966
 Db 3304 TTGAAATGACCAAGCTTAATGAGGTGCACTTAAGCTTAATTAATTAATTAATTAAT 3363
 QY 2967 GAATTCGGCTTCATTAACCATTAATAAGAGAGATTAAGCAATGATTAATTAATTAAT 3026
 Db 3364 GAATTCGGCTTCATTAACCATTAATAAGAGAGATTAAGCAATGATTAATTAATTAAT 3423
 QY 3027 GAGCAGACGAGACGAAATTAAGAGCAAAAGTTGACCTGCTTAATTAATTAATTAAT 3086
 Db 3424 GAGCAGACGAGACGAAATTAAGAGCAAAAGTTGACCTGCTTAATTAATTAATTAAT 3483
 QY 3087 GGTGAGCGCAAAAGTGGCTCAAGAGAGAGG----- 3119
 Db 3484 AGTAATGCAAGAGTGGCTCAAAAGAGCGGTTTCTGATTAACCTGCTTAATTAAT 3543
 QY 3120 ----- 3119
 Db 3544 CAGTTAGAGTATTAACAGCGCAACAAAGTTGACCTGCTTAATTAATTAATTAAT 3603
 QY 3120 -----AGAGCAGGTTTCTGATTAACCTGCTTAATTAATTAATTAATTAAT 3158
 Db 3604 GCAAAAAGAGTGGCTCAAAAGAGCGGTTTCTGATTAACCTGCTTAATTAATTAAT 3663
 QY 3159 TTAAGCGCTTAAGAGCCAAACAAAGCTGAACCTGCTTAATTAATTAATTAATTAAT 3218
 Db 3664 TTAAGCGCTTAAGAGCCAAACAAAGCTGAACCTGCTTAATTAATTAATTAATTAAT 3273
 QY 3219 AAAAAAAGAGTGGCTCAAAAGAGCGGTTTCTGATTAACCTGCTTAATTAATTAAT 3278
 Db 3724 AAAAAAGTGGCTCAAAAGAGCGGAGAGGTTTCTGATTAACCTGCTTAATTAATTAAT 3783
 QY 3279 CTGTTT---GCAATTAAGAGCCGCACTTAAGGTTTATTAAGCCCAACAGCAATGAGAAA 3335
 Db 3784 CGGTTAAAGGTTATGAAGTCAAACTTAAGGTTTATTAAGCCCAACAGCAATGAGAAA 3843
 QY 3336 GATGCTGCTCAAG 3395
 Db 3844 GAACTCAAGATCA-----GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3894
 QY 3396 TATTCAAATTAAGTGGCTTAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3455
 Db 3896 TATTCAAATTAAGTGGCTTAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3954
 QY 3456 GATGAATTAAGTGGCTTTTGTGATTAAGAGCAAACTGCGGTTGAGCAAAATTAAGCA 3515
 Db 3956 GATGAATTAAGTGGCTTTTGTGATTAAGAGCAAACTGCGGTTGAGCAAAATTAAGCA 4014

QY 3516 CAGATTAAG 3575
 Db 4015 CAGATTAAG 4071
 QY 3576 TTAAGTCAAAATTAAGGAGTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3635
 Db 4072 TTAAGTCAAAATTAAGGAGTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4131
 QY 3636 CATACCGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3695
 Db 4132 CATACCGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4191
 QY 3696 ATGTGAGGTTTGGCCCATATCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3755
 Db 4192 ATGTGAGGTTTGGCCCATATCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4251
 QY 3756 GGAATCAAGTGGAGTAAATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3815
 Db 4252 GGAATCAAGTGGAGTAAATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4311
 QY 3816 AATTAAGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 3875
 Db 4312 AATTAAGGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 4371
 QY 3876 TTTGAGGTTAATGCTAATTTATTAAGAGTGAATGAATGAATGAATGAATGAATGAATGAAT 3935
 Db 4372 TTTGAGGTTAATGCTAATTTATTAAGAGTGAATGAATGAATGAATGAATGAATGAATGAAT 4431
 QY 3936 AAAAAAGCTTACCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3995
 Db 4432 AAAAAAGCTTACCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4491
 QY 3996 ACTCGAGAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4055
 Db 4492 ACTCGAGAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4551
 QY 4056 AACGCTAAGTGAACCAACAGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4115
 Db 4552 AACGCTAAGTGAACCAACAGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4611
 QY 4116 TGGCAAAAGAGAGTGGAGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4175
 Db 4612 TGGCAAAAGAGAGTGGAGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4671
 QY 4176 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4235
 Db 4672 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4731
 QY 4236 CGTGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4295
 Db 4732 CGTGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4788
 QY 4296 TCCCACTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4319
 Db 4789 TCCCACTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4810
 RESULT 6
 ADA09357
 ID ADA09357 standard; DNA, 4828 BP.
 XX
 AC ADA09357:
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Haemophilus HAP protein gene from strain 1396B.
 XX
 KW de; Haemophilus adherence and penetration protein; HAP; gene;
 KW anti-bacterial; immunostimulant; vaccine; Haemophilus influenza infection;
 XX influenza.
 XX
 OS Haemophilus influenzae; strain 1396B.
 XX

US2003073166-A1.
 17-APR-2003.
 22-FEB-2002; 200205-00080505.
 25-AUG-1994; 94US-00296791.
 20-APR-2001; 2001US-00839996.
 (GENE/) GENE J W S.
 Gene JMS;
 WPI: 2003-56738/53.
 P-PSDB: ADA09358.
 New recombinant Haemophilus adhesion and penetration protein, useful as a vaccine for prophylactic or therapeutic treatment of a Haemophilus influenzae infections, in generating antibodies for passive immunization.
 Claim 1; Fig 24; 149p; English.

The invention relates to a new recombinant Haemophilus adhesion and penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351, ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic acid that will hybridize under stringency conditions to the nucleic acid detailed above, an expression vector comprising transcriptional and translation regulatory nucleic acid operably linked to the HAP nucleic acid, producing HAP by culturing a host cell transformed with the expression vector, a composition comprising a pharmaceutical carrier and HAP, inducing an immune response in a patient by administering the composition, a composition comprising an antibody capable of binding to HAP (or to a peptide selected from 35 peptide sequences appearing as ADA09353-ADA09357), and an antibody capable of binding to the peptides. The Haemophilus adhesion and penetration (HAP) protein is useful as a vaccine for prophylactic or therapeutic treatment of an H. influenzae infection (e.g. influenza), in generating an immune response in a patient, and in generating antibodies to the HAP protein. The antibodies may be used to diagnose the presence of H. influenzae infection in a sample or patient, in passive immunisation treatments, and to purify or separate HAP proteins or the H. influenzae organism from a sample. The present sequence is a HAP gene from a strain of H. influenzae.

Sequence 4828 BP; 1641 A; 896 C; 993 G; 1298 T; 0 U; 0 Other;
 Query Match 68.8%; Score 2970; DB 8; Length 4828;
 Best Local Similarity 82.3%; Freq. No. 0;
 Matches 3581; Conservative 0; Mismatches 645; Indels 124; Gaps 10;

1 TCATAGTCTTACTA-GTATTTTAAACGAAATTAATTAATTAATTAACAT 59
 253 TCATAGTCTTAAACGATTTTAAACGAAATTAATTAATTAATTAACAT 312
 60 ATGAAAAAACTGATTTGCTTAATTTTAAACGCTGATTTCAATAGGATAGTA 119
 313 ATGAAAAAACTGATTTGCTTAATTTTAAACGCTGATTTCAATAGGATAGTA 372
 120 TCGAAGCGTGGCTGTCACATTTTGGATGATATCCAAATTAATTAATTAAT 179
 373 TCGAAGCGTGGCTGTCACATTTTGGATGATATCCAAATTAATTAATTAAT 432
 180 GCCGGAATTAAGGAGTTCACAGTTGGGCTCAAAATTAATTAATTAATTAAT 239
 433 GCCGGAATTAAGGAGTTCACAGTTGGGCTCAAAATTAATTAATTAATTAAT 492
 240 GGGCAATTAAGTGGCAATCAATGACAAAGCCCGATGATTTTCTGATGTC 299
 493 GGAATTAAGTGGCAATCAATGACAAAGCCCGATGATTTTCTGATGTC 552
 300 GGTAAAGGCTGGCAATCAATGACAAAGCCCGATGATTTTCTGATGTC 359
 553 GGAATTAAGGCTGGCAATCAATGACAAAGCCCGATGATTTTCTGATGTC 612

360 GATATACAGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 419
 613 GGTATACAGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 672
 420 TATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 479
 673 TATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 732
 480 TATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 539
 733 TATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 792
 540 TATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 599
 793 TATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 852
 600 TCGAAGCGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 659
 853 TCGAAGCGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 912
 660 CATTAATTAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 719
 913 CATTAATTAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 972
 720 TCGAAGCGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 779
 973 TCGAAGCGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1032
 780 GGGCAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 839
 1033 GGGCAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1092
 840 GGGCAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 899
 1093 GGGCAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1152
 903 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 959
 1153 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1212
 960 GGTAAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1019
 1213 GGTAAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1272
 1020 TCGAAGCGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1079
 1273 TCGAAGCGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1332
 1080 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1139
 1333 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1386
 1140 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1195
 1387 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1446
 1200 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1259
 1447 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1506
 1260 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1319
 1507 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1566
 1320 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1379
 1567 AATTAATTTGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 1626
 1380 GGTAAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1439
 1627 GGTAAAGGCTTTTGGCGAAATTAATCAAGCAAGGCGCAAGTTGCCGTGATAT 1686

QY 1440 CAGGAGAGATCAAGGCAACAAAGCCTTTAGTGAATGGCTGCTGAGGCGAG 1499
 DB 1687 CAGGAGAGATCAAGGCAACAAAGCCTTTAGTGAATGGCTGCTGAGGCGAG 1746
 QY 1500 GGGACCTGCTAATTAAGATGATTAACAAATTTGATACCGATTAATTTATTTGGGCTT 1559
 DB 1747 GGTACGGTTCAGTTAAAGATGACAAAGCAATTTAAATATATTAATTTATTTGGGCTT 1806
 QY 1560 CGTGGTGGCTGCTGATCTTAACGGGCAATTCATTAACCTTTAAAGATTCGAAATAG 1619
 DB 1867 CGTGGTGGCTGCTGATCTTAACGGGCAATTCATTAACCTTTAAAGATTCGAAATAG 1866
 QY 1620 GAGGAGGAGGCAATGATTTGAAACATTAATTAACCTTAACGGGCAATTCATTAAC 1679
 DB 1867 GATGAGGAGGCAATGATTTGAAACATTAATTAACCTTAACGGGCAATTCATTAAC 1926
 QY 1660 GGGAGGAGGCAATGATTTGAAACATTAATTAACCTTAACGGGCAATTCATTAAC 1739
 DB 1927 GGGAGGAGGCAATGATTTGAAACATTAATTAACCTTAACGGGCAATTCATTAAC 1977
 QY 1740 GAAATGGCTCAACAGGCTGCTTGGCGAAACAGTAAATTAACATTAACGATTCG 1799
 DB 1978 ---ATTCCTTTAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2034
 QY 1800 AACCTTATTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1859
 DB 2035 AATGCTTCAATCAATCCGCTTAAACAAAGATTAATCACTTCCCTTCAAGGAGGCAAT 2094
 QY 1860 TTAAGGAGGATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1919
 DB 2095 TTAAGGAGGATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2154
 QY 1920 CAGGCTCAACATCATTTAAACGTTGTCAGAAATGGAAGTAAACCAACCAACCAAC 1979
 DB 2155 CAGGCTCAACATCATTTAAACGTTGTCAGAAATGGAAGTAAACCAACCAACCAAC 2208
 QY 1980 ATGCTGAGATCAAGATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2039
 DB 2209 GTGGTATTTGATGAGATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2268
 QY 2040 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2099
 DB 2269 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2328
 QY 2100 AATTAAGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2159
 DB 2329 AATTAAGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2388
 QY 2160 TCAATTTGCAAGATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2219
 DB 2389 TCAATTTGCAAGATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2448
 QY 2220 AATCTATTCGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2279
 DB 2449 GATTCATTCGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2508
 QY 2280 AATCTATTCGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2339
 DB 2509 AATCTATTCGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2568
 QY 2340 AATCTATTCGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2399
 DB 2569 AATCTATTCGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2628
 QY 2400 AATCTATTCGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2459
 DB 2629 AATCTATTCGCAAAATGCAATTTGCTGTCGCAAAATGCAAAATGCAAAATGCAAAATG 2688
 QY 2460 TTAAGGAGGATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2519
 DB 2689 TTAAGGAGGATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2748
 QY 2520 GAAATGGCTCAACAGGCTGCTTGGCGAAACAGTAAATTAACGTTAAATTAAC 2579

DB 2749 GAAATGGCTCAACAGGCTGCTTGGCGAAACAGTAAATTAACGTTAAATTAAC 2808
 QY 2580 AGTACGATCAAGTAAATTAACGATTAACGATTAACGATTAACGATTAACGATTAAC 2636
 DB 2809 AGTACGATCAAGTAAATTAACGATTAACGATTAACGATTAACGATTAACGATTAAC 2868
 QY 2637 CGTTCATTAAGAGCGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2696
 DB 2869 CGTTCATTAAGAGCGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2928
 QY 2697 GTTAATGCTAAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2756
 DB 2929 GTTAATGCTAAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2988
 QY 2757 AAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2816
 DB 2989 AAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3048
 QY 2817 AACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2876
 DB 3049 AACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3108
 QY 2877 CAGGCTTATCAAGTAAATTAACGATTAACGATTAACGATTAACGATTAACGATTAAC 2936
 DB 3109 CAGGCTTATCAAGTAAATTAACGATTAACGATTAACGATTAACGATTAACGATTAAC 3168
 QY 2937 TTAAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2996
 DB 3169 TTAAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3228
 QY 2997 CAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3056
 DB 3229 CAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3288
 QY 3057 CAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3116
 DB 3289 CAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3342
 QY 3117 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3176
 DB 3343 ---AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3399
 QY 3177 AAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3231
 DB 3400 AAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3459
 QY 3232 TGGGCTCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3287
 DB 3460 TGGGCTCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3519
 QY 3288 TTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3321
 DB 3520 TTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3579
 QY 3322 ---AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3332
 DB 3580 GCTTCATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3639
 QY 3333 AAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3392
 DB 3640 AAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3699
 QY 3393 CGTTCATTAAGAGCGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 3452
 DB 3703 CGTTCATTAAGAGCGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 3759
 QY 3453 CAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3512
 DB 3760 CAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3819
 QY 3513 GAGGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3572

```

Db 3820 TCACAGGATTAAGAGCTTATGATTCGATCGCTTCCTGTTT---CAGCAGAAAACG 3876
Qy 3573 AACTACGTCMAATTG93GATGCAAAAAGCCTTAGCTAAATGAGCAATG93GAGTTTC 3632
Db 3877 AACTGCTCAATTG93GATGCAAAAAGCCTTAGCTAAATGAGCAATG93GAGTTTC 3936
Qy 3633 TCAGATAGCCGTTGATGATTAATCCTTTGATGAACAGGTTAAATATCAACCGCATTAACG 3692
Db 3937 TCAGATAGCCGTTGATGATTAATCCTTTGATGAACAGGTTAAATATCAACCGCATTAACG 3996
Qy 3693 ATGATGTCGGGTTTGGCCCATATCATG93GCGATTTACAGATTTG93GTAAGT93GA 3752
Db 3997 ATGATGTCGGGTTTGGCCCATATCATG93GCGATTTACAGATTTG93GTAAGT93GA 4056
Qy 3753 ACGGGAATCAGTGCAGATTAATGCTGAAAGCAACACCGAAAATTCATCGAAAACG 3812
Db 4057 ACGGGAATTAAGTGCAGATTAATGCTGAAAGCAACACCGAAAATTCATCGAAAACG 4116
Qy 3813 ATAAATTATG93GATGCAATGCAATGATTCAGTTCCGTTTAA93GCAATTCAGCT 3872
Db 4117 ATAAATTATG93GATGCAATGCAATGATTCAGTTCCGTTTAA93GCAATTCAGCT 4176
Qy 3873 TATTTGAGATTAATGCTATTTATTTATGAAAGGTAATATCATCTGAGAGT93GA 3932
Db 4177 TATTTGAGATTAATGCTATTTATTTATGAAAGGTAATATCATCTGAGAGT93GA 4236
Qy 3933 GTGAAAACGCTTACGCTTCAATTAATGCTAATTAATGCTGCAATTCAGTTGATTAACA 3992
Db 4237 GTGAAAACGCTTACGCTTCAATTAATGCTAATTAATGCTGCAATTCAGTTGATTAACA 4296
Qy 3993 TTTTCTCCAGCATATATGAGGTTAAAGCCTTATTTCTTCTGCAATTTATGATGTT 4052
Db 4297 TTTTCTCCAGCATATATGAGGTTAAAGCCTTATTTCTTCTGCAATTTATGATGTT 4356
Qy 4053 TCAGAGCCTTACGCTAATTAATGCTAATTAATGCTGCAATTCAGTTGATTAACA 4112
Db 4357 TCAGAGCCTTACGCTAATTAATGCTAATTAATGCTGCAATTCAGTTGATTAACA 4416
Qy 4113 TATTTGCAAAAAGAGT93GATTAAGCGAATTTTATTCATTTCCAAATTCGCTTTT 4172
Db 4417 TATTTGCAAAAAGAGT93GATTAAGCGAATTTTATTCATTTCCAAATTCGCTTTT 4476
Qy 4173 ATTCGAAATTCAGGTTTCAGATTCGCTGCAATTAAGGTAATG93GCTGAAATG93G 4232
Db 4477 ATTCGAAATTCAGGTTTCAGATTCGCTGCAATTAAGGTAATG93GCTGAAATG93G 4536
Qy 4233 TATCGTTGTAATAATCAATTAATTTAT 4262
Db 4537 TATCGTTGTAATAATCAATTAATTTAT 4566

```

```

PR 25-AUG-1994: 94US-00236791.
PR 20-APR-2001: 2001US-00839996.
PA (GENE//) GENE J M S.
PI Gene JMS;
XX MPI: 2003-567308/53.
XX P-PSDB: ADA09356.
DR
DR
PT New recombinant Haemophilus adherence and penetration protein, useful as a
PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
PT influenzae infections, in generating antibodies for passive immunization.
XX
XX Claim 1; Fig 22; 149pp; Eng.-lit.
PS
CC The invention relates to a new recombinant Haemophilus adherence and
CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
CC ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
CC acid that will hybridize under stringency conditions to the nucleic acid
CC detailed above, an expression vector comprising transcriptional and
CC translation regulatory nucleic acid operably linked to the HAP nucleic
CC acid, producing HAP) by culturing a host cell transformed with the
CC expression vector, a composition comprising a pharmaceutical carrier and
CC HAP, inducing an immune response in a patient by administering the
CC composition, a composition comprising an antibody capable of binding to
CC HAP (or to a peptide selected from 35 peptide sequences appearing as
CC ADA09359-ADA09393), and an antibody capable of binding to the peptides.
CC The Haemophilus adherence and penetration (HAP) protein is useful as a
CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
CC infection (e.g., influenza), in generating an immune response in a
CC patient, and in generating antibodies to the HAP protein. The antibodies
CC may be used to diagnose the presence of H. influenzae infection in a
CC sample or patient, in passive immunisation treatments, and to purify or
CC separate HAP proteins or the H. influenzae organism from a sample. The
CC present sequence is a HAP gene from a strain of H. influenzae.
XX
XX Sequence 4822 BP; 1652 A; 818 C; 982 G; 1370 T; 0 U; 0 Other:
SQ
Query Match 54.1%; Score 2770.2; DB 8; Length 4822;
Best Local Similarity 73.6%; Pred. No. 0;
Matches 3466; Conservative 0; Mismatches 813; Indels 76; Gaps 13:
Qy 1 TCATATGTCGTTTACTA-GTATTTTATATACGAAATTTCTTAATTAATTAACATT 59
Db 328 TCATATGTCGTTTACTA-CGATATTTTATATACGAAATTTCTTAATTAATTAACATT 387
Qy 60 ATGAAAAAACTGATATTTGCTTAAATTTTAAACGCTGATTTGATGATGATGATGAT 119
Db 388 ATGAAAAAACTGATATTTGCTTAAATTTTAAACGCTGATTTGATGATGATGATGAT 447
Qy 123 TCAGAGCCTG93GCTGCTACACTATTTTGGATGATTAACAATA-TAATGATTTT 179
Db 448 TCAGAGCCTG93GCTGCTACACTATTTTGGATGATTAACAATA-TAATGATTTT 507
Qy 180 GCCGAGATTAAGGAAAGTTTACAGTTT93GCTGCTGAAATTTAAGTTTATACAAACA 239
Db 508 GCCGAGATTAAGGAAAGTTTACAGTTT93GCTGCTGAAATTTAAGTTTATACAAACA 567
Qy 240 GGGCATATGTTGCAACATTAATGCAAAAGCCCGATGATTTTTCGTAAGTGTCA 299
Db 568 GGGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
Qy 300 CGTACGCGTGGAGGCTTGTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 359
Db 628 CGTACGCGTGGAGGCTTGTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 687
Qy 360 GATATATGATGATTTTGTGTAAGGAAACACCCGATGATGATGATGATGATGATGAT 419
Db 688 GATATATGATGATTTTGTGTAAGGAAACACCCGATGATGATGATGATGATGATGATGAT 747
Qy 420 TATTAAGTGTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 479

```

```

RESULT 7
ADA09355
ID ADA09355 standard; DNA; 4822 BP.
XX
XX ADA09355;
XX
XX 06-NOV-2003 (first entry)
XX
XX Haemophilus HAP protein gene from strain 3219B.
XX
XX db; Haemophilus adherence and penetration protein; HAP; gene;
XX anti-bacterial; immunostimulant; vaccine; Haemophilus influenzae infection;
XX influenza.
XX
XX Haemophilus influenzae; strain 3219B.
XX
XX US003073166-A1.
XX
XX 17-APR-2003.
XX
XX 22-FEB-2002; 2002US-00080505.
XX

```


000	001	002	003	004	005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035	036	037	038	039	040	041	042	043	044	045	046	047	048	049	050	051	052	053	054	055	056	057	058	059	060	061	062	063	064	065	066	067	068	069	070	071	072	073	074	075	076	077	078	079	080	081	082	083	084	085	086	087	088	089	090	091	092	093	094	095
000	001	002	003	004	005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035	036	037	038	039	040	041	042	043	044	045	046	047	048	049	050	051	052	053	054	055	056	057	058	059	060	061	062	063	064	065	066	067	068	069	070	071	072	073	074	075	076	077	078	079	080	081	082	083	084	085	086	087	088	089	090	091	092	093	094	095

Qy	540	AAATGAAATGGCGAGTCACTTATTCAGATGAGACAAATAATCCAAAGCGGTTCGATCGCC	599
Db	986	TAATGGAATGGTAACTATTCAAAAATTTTAACTCAATATCTTTGGCGATTAAGATTGGA	955
Qy	600	CTGACAGCGCGAGTGTTCGGCAATGATCAAGACAAAGGC---GACCAAGTTGCCCGGTGCA	656
Db	956	AGTGGGCATCAATGGTGGAAAGACGATTAATATTAACCAATTGGAGACTTGGACCTTAGGA	1015
Qy	657	TATCATATCTGACAGCTGGGCAATNCACACATCGCGTGGAGCGAGTAATGGAATATGCG	716
Db	1016	GGTTCAATGGTTAAATAGTGGAAATACCTTTGAGATGGACACAGCTGATGACGTACATTA	1075
Qy	717	TATTTGGAGGCGCAGTGGTGTGTAAGCGGAGAAATATGGTCCATTACCATAGCGAGCTCA	776
Db	1076	GAAATTAAATGGCGAGTACGAAATATCTAAATAAATATGCTCATACATACGCGAGGTTGA	1135
Qy	777	AAGGCGGAGCAGTGGTCTCCGATGTTTATTATATGATGCTGAAAACAAATAGGTAAAT	836
Db	1136	TTGCGGGGCAATGGTGTCTCCGATGTTTATTATATGATGCTGAAAACAAATAGGTAAATTA	1195
Qy	837	AATGGCAATATATACGGGCAAGGCAACCCCTTTGAAAGGCAAGAAATAGGGTTCAATGGT	896
Db	1196	AATGGCGGTGTACGTGAGGAAATCTTATGCTGCAGATGAGAAACGCTATCAATATTA	1255
Qy	897	CGCAATCTTATTTATGTAATATTTTGGAAAGATTTACATCATCATCTTACACCGCA	956
Db	1256	CGAAAGATATTTTTCAGGTATTTCTTAATCAAGACATCACTACACTAATTTTGGGATACT	1315
Qy	957	GCCTGTAATGAGAGTNCACATTAATGAGAAATGATATGCTAGGCGCTC-----	1006
Db	1316	AATGCGAATATACATTTATATATGAGGATGACCAATGAGAGATGGCAACAAATCAA	1375
Qy	1097	--TATTACTCGAATACAGAAATACCATCAAGAAATTAATAATTCGTTAGCAATATAGCT	1064
Db	1376	AGTACATTAACCTAAAAACCTATTAACCTGAAAGCAATAGTGGCTTTATGATTAATAC	1435
Qy	1065	TTACCTTTGAAAAGAAAGATTAA-----GTTCAATCTTGATATGAGACCTAAT	1118
Db	1436	CAACTTCATGATGCTGAGAAATAAATGCGATGATATCCCTTTAATAAGGTCTAAT	1495
Qy	1119	ATTATATTCACAGTTTAAACATGAGAGAAAGGATATTTTATGAGATCAAAAACAAGA	1178
Db	1486	CGATGTCGCAAGTATTAATCAATGAGAAAGATATTACTTTGGCATACAGAAACAAGA	1555
Qy	1179	TCATTAATCTTCGATCTGACATTAACCAAGGCGGTGCTTTATTTTGAAGCGTAAAT	1238
Db	1556	ACTTAAACATTTGAAATATATATTAATCAAGGTGCGAGGTGATATGATTTTGAAGCGTAAAT	1615
Qy	1239	TTTACGATCTCCAAAT---TCTAACCAACTGCGCAAGGACTGCGCATCATGTAAGT	1295
Db	1616	TTTGTTGTAAGGCAATCAAAATTAATTAATCTGGCAAGGTGCAGAGCTTTCTGTTGGA	1675
Qy	1296	GAAATAGCACCGTACTCTTGAAAGTAAATGCGGTGGAACATGATCGACTTTCTAAAT	1355
Db	1676	GAGAAAGTACTGTGATGATGAGAGGTGCATTAATCCAAAGAGCGATCGCTTACCAAAAT	1735
Qy	1356	GGTAAAGGAACTATGCACTTCAACCCAAAGGGAATAAAGTTGCATGACGCTAGGC	1415
Db	1736	GGCGTCGGAACTTACTCTGTAAATGTTAAAGGAAAAAATCTAGGAAGCTTAAGTGGGT	1795
Qy	1416	GATGCTAAAGTCAATTTTGGAGAGACAGCGACATCAAGGCAACAAACAGACTTAAAT	1475
Db	1796	AACGCTTTGTTGTTATATCAACAGCAATGATACGGTCAAAAACAGCCTTAAAT	1855
Qy	1476	GAAATGCTGTTAGCGGCAAGGGAATGTTCAATTAACGATGATTAACAAATTAAT	1535
Db	1856	GAGTGGCATTTGAATGATGATGAGCTACCGTTCACTAAATATGTCAGATCAAGTTGAT	1915
Qy	1536	ACCGAATAAATTTATTTGGCTTCGATGCGGTTCCTAGACTTTAAAGGCAATCATTA	1595
Db	1916	CTTAACGATTAATTTGCGCTTCGATGCGGTTCCTTAAGTCTTAAATGAGCAATCATTA	1975

QY 1596 ACCTTAAACGATATCAAAATATGCGACGAGGCGCATGATTTGTGACATTAATCAACT 1655
 DB 1976 ACCTTAAACGATATCAAAATATGCGAGGAGGCGATGATTTGTGACATCAACAGCTCT 2035
 QY 1656 CAAGCCGCTAATGTCATTAATCTGCGACGAAAGCATTTGTTCTACTAATGGAATAT 1715
 DB 2036 CAAGCCGCTAATGTCATTAATCTGCGACGAAAGCATTAATCAACAGCTCTCT 2083
 QY 1716 ATTAATTAATCTGATTAAGAAAGAAATGCTCAAGAGTTGTTGCGAAACGAT 1775
 DB 2084 AGCAACACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2143
 QY 1776 AAAAATTAACACATGCGCATTAACCTTAATTAATTAATTAATTAATTAATTAAT 1835
 DB 2144 AAAGCTAAACAAATGCTGTTTAAATGTAATTAATTAATTAATTAATTAATTAAT 2203
 QY 1836 TTGCTACTTCAAGTGTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1895
 DB 2204 TTGCTGCTTCTGCGGGAACAAATTAATTAATTAATTAATTAATTAATTAATTA 2263
 QY 1896 TTTTCAGCGGTAGACGACGCGACGCTCAATCAATTAATTAATTAATTAATTAAT 1955
 DB 2264 GTTTTATGCTGTGTCACACCTCAATCAATTAATTAATTAATTAATTAATTAAT 2323
 QY 1956 ATGAGAGTATACCAACAGGCAATTTGTGTGATCAAGATTAATTAATTAATTAAT 2015
 DB 2324 ATGAGAGTATACCAACAGGCAATTTGTGTGATCAAGATTAATTAATTAATTAAT 2383
 QY 2016 AAAGCTAAACCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2075
 DB 2384 AAAGCTAAACCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2443
 QY 2076 ATGAGAGGAAATTTGACAGTCAAGATTAATTAATTAATTAATTAATTAATTAAT 2135
 DB 2444 ATGAGAGGAAATTTGACAGTCAAGATTAATTAATTAATTAATTAATTAATTAAT 2503
 QY 2136 CAACAAATATCAATTTGACAGCTCAATTAATTAATTAATTAATTAATTAATTAAT 2195
 DB 2504 CAACAAATATCAATTTGACAGCTCAATTAATTAATTAATTAATTAATTAATTAAT 2563
 QY 2196 GATTTAATCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2255
 DB 2564 GATTTAATCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2623
 QY 2256 AATTTAATCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2315
 DB 2624 AATTTAATCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2683
 QY 2316 ACTTTAATCAATCAAGCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2375
 DB 2684 ACTTTAATCAATCAAGCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2743
 QY 2376 GATTTTCCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2435
 DB 2744 AATCTTTCAATCAAGCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2803
 QY 2436 TTAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2495
 DB 2804 TTAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2863
 QY 2496 GACAAAGCAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2555
 DB 2864 GACAAAGCAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2923
 QY 2556 TTGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2615
 DB 2924 TTGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2983
 QY 2616 TCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2669
 DB 2984 TCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3043
 QY 2670 GCGAAGATGCTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2729

DB 3044 GCGAAGATGCTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3103
 QY 2730 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2789
 DB 3104 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3163
 QY 2790 GAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2849
 DB 3164 GAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3223
 QY 2850 TTAACCTTGTGTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2909
 DB 3224 TTAACCTTGTGTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3283
 QY 2910 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2969
 DB 3284 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3343
 QY 2970 TTCCGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3029
 DB 3344 TTCCGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3403
 QY 3330 CAAGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3089
 DB 3404 CAAGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3463
 QY 3090 GAGCCAAAGTGGGTGAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3149
 DB 3464 GAGCCAAAGTGGGTGAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3514
 QY 3150 CAAGCTGTTTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3209
 DB 3515 CAAGCTGTTTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3571
 QY 3210 AGTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3269
 DB 3572 AGTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3631
 QY 3270 GATCAAGCTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3329
 DB 3632 GATCAAGCTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3682
 QY 3330 GAAAGATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3386
 DB 3683 GAAAGATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3742
 QY 3387 ATCAAGCTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3446
 DB 3743 ATCAAGCTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3802
 QY 3447 TCGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3506
 DB 3803 TCGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3862
 QY 3507 AATATGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3566
 DB 3863 AATATGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3923
 QY 3567 AAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3626
 DB 3924 AAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3979
 QY 3627 GTTTCTGCAATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3686
 DB 3980 GTTTCTGCAATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4039
 QY 3687 TTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3746
 DB 4040 TTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4099
 QY 3747 GTGGAAGCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3806

Db 4100 GTGGGTGGGGAATTGCTGAGTAATAATGCTGAGAAACAAAGCCGAAATTCATCGA 4159
 Qy 3807 AAGCGATAAATTAATGCGTGAATGCAAGTATCATGTCGCTTAAGGCAATGGGCAAT 3866
 Db 4160 AAGCGATAAATTAATGCGTGAATGCAAGTATCATGTCGCTTAAGGCAATGGGCAAT 4219
 Qy 3867 CAGCCTTAATTTGAGAGTTAATCGTAATTTTAATGACGTAATAATTAATCAATCTGAGGAA 3926
 Db 4220 CAGCCTTAATTTGAGAGTTAATCGTAATTTTAATGACGTAATAATTAATCAATCTGAGGAA 4279
 Qy 3927 GTGAGAGTGAAGAAAGCCTACCTGCTGATTAATGCTTAATGCTGATTAATGCTGATTAAT 3986
 Db 4280 GTGAGAGTGAAGAAAGCCTACCTGCTGATTAATGCTTAATGCTGATTAATGCTGATTAAT 4339
 Qy 3987 TATACATTTATCCCGAGAGATTAATATGAGCTTAATGCTTAATGCTGATTAATGCT 4046
 Db 4340 TATACATTTATCCCGAGAGATTAATATGAGCTTAATGCTTAATGCTGATTAATGCT 4399
 Qy 4047 GATGTTTCAAAAGCTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4106
 Db 4400 GATGTTTCAAAAGCTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4459
 Qy 4107 GAGCTTAATGCGCAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4166
 Db 4460 GAGCTTAATGCGCAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4519
 Qy 4167 GCTTTATCTCAAAATCTCAAGGTTTCACTGCGGCAACAGCAAAATGTTGGGCGTGAAG 4226
 Db 4520 GCTTTATCTCAAAATCTCAAGGTTTCACTGCGGCAACAGCAAAATGTTGGGCGTGAAG 4579
 Qy 4227 TTGGGCTATCGTTGGTAAATCAAC 4252
 Db 4580 TTGGGCTATCGTTGGTAAATCAAC 4605

RESULT 9

ACA41657
 ID ACA41657 standard; DNA; 4350 BP.

XX AC ACA41657;
 XX AC ACA41657;
 Df 19-JUN-2003 (first entry)
 XX Prokaryotic essential gene #23314.
 DE Prokaryotic essential gene #23314.
 XX
 KM Antisense; ds; prokaryotic essential gene; cell proliferation;
 XX drug design; gene.
 XX
 OS Neisseria meningitidis.
 XX
 PN W020027183-A2.
 XX
 PD 03-OCT-2002.
 PF 21-MAR-2002; 2002MO-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00672851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KJ, Zyskind JW,
 PI Wall D, Trivick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HJ.
 XX
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU37787.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX
 PS Claim 14; SEQ ID NO 29527; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound an antibiotic; (10) profiling a
 CC organism's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 4350 BP; 1237 A; 1186 C; 1072 G; 855 T; 0 U; 0 Other;

Query Match 31.3%; Score 1353.8; DB 7; Length 4350;
 Best Local Similarity 59.7%; Pred. No. 0;
 Matches 2586; Conservative 0; Mismatches 1527; Indels 216; Gaps 10;

Qy 81 CTTAATTTTAAACCTGCTTCAATTTAGGATGATGCGAAGCGTGGCTGCTAC 140
 Db 73 CCTGCTTACTTACCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132
 Qy 141 ACTTAATTTTAAACCTGCTTCAATTTAGGATGATGCGAAGCGTGGCTGCTAC 200
 Db 133 ACTTAATTTTAAACCTGCTTCAATTTAGGATGATGCGAAGCGTGGCTGCTAC 192
 Qy 201 ACGTGGGCTCAAAATTAAGTTTATTAACAAACAGGCGAATAGTTGGCACATCA 260
 Db 193 GCGTGGGCTCAAAATTAAGTTTATTAACAAACAGGCGAATAGTTGGCACATCA 252
 Qy 261 ATGCAAAAGCCCGATGATTTATTTTCTGATGATGATGATGATGATGATGATGATGATGAT 320
 Db 253 ATGCAAAAGCCCGATGATTTATTTTCTGATGATGATGATGATGATGATGATGATGATGAT 312
 Qy 321 GTTAAATTAATTAATTTGAGCGTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 383
 Db 313 GTTAAATTAATTAATTTGAGCGTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 372
 Qy 381 GTTAAATTAATTAATTTGAGCGTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 443
 Db 373 GTTAAATTAATTAATTTGAGCGTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 432
 Qy 441 AATTAATTAATTAATTTGAGCGTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 500
 Db 433 AATTAATTAATTAATTTGAGCGTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 492
 Qy 501 AATTAATTAATTAATTTGAGCGTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 560
 Db 493 AATTAATTAATTAATTTGAGCGTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 552

QY 561 TCAGATAGACAAATATCCAGAACGTTGCTGATCCGCTCTGGACGGAGTTTGGCGA 620
 DB 553 TCGATATAGAAAAATATCCGACGCTGCCATCCGCTTAGGACACACATATGGGCT 612
 QY 621 AATGATAGACAAAGGCGACCAAGTTCCGCTGATATCATATCTGACAGCTGGCAAT 680
 DB 613 TATGATGATGACAAACGCGCATTTATCTCTCCGCGCATGTTAATTGGCGCAT 672
 QY 681 ACACAGATCAGCGTGGGACGATGATGATGATGATGATGATGATGATGATGATGAT 740
 DB 673 ACACATATGAGGGGTTGGGAAATATATGAGTATGATGATGATGATGATGATGAT 732
 QY 741 GCGGAGAAATATGCTCCATATCCGATGACAGCTCAAGGGGACAGTGTCTCCGATG 800
 DB 733 GCGACGACCTATGCGCTTATCCGATTCAGGTTGGGACGACGACGCTTCCGATG 792
 QY 801 TTTATTTATGATGCTGAAAAAAGAAAAATGTTATTTATTTATTTATTTATTTATTTAT 860
 DB 793 TTTATTTATGACAAAAAAGAAAAATGATTTATTTATTTATTTATTTATTTATTTAT 852
 QY 861 CCTTTGAGGCAAGAAAAATGGTTTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 917
 DB 853 CCTTATTCGGGAGGAAAAAGGTTTCCAGCTGATAGCAAGTATGTTTCTAGATGAC 912
 QY 918 ATTTTGAAGAAATTTACATCATCACTTTACACCGAGCTGATATGATGATGATGATGAT 977
 DB 913 ATTTTGAAGAGGAAATCATCATACCGCTTTTGAACCGGCGATGATGATGATGATGAT 972
 QY 978 ATTTGTAAGAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1037
 DB 973 TTTACATCCAAACAAACAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1032
 QY 1038 ATTTAAATTTACGTTAGCAATATGATGTTT-----ACCTTTGAAGAGAGGATTA 1088
 DB 1033 CCAAGCTTTAAAGTACAGACAGCTCCGACTTTGACAGATTTTGAATTAATCTGATTA 1092
 QY 1089 GTTCATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148
 DB 1093 GAACCACTTTACGCGGCGAGGGGGTGTAAATCAGTACCGTCAAGGTTAAACAGGTTAA 1152
 QY 1149 ACGCTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1208
 DB 1153 AACCTTCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
 QY 1209 GGGCGGGGTTGCTTTATTTTGAAGGTTATTTTACATGATCTCAAACTCTCAACAACT 1268
 DB 1213 GGGCGGGGCTTTGATTTTGAAGGTTATTTTACGCTCTCGCTGATGATGATGATGATGAT 1272
 QY 1269 TGGCAAGAGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1328
 DB 1273 TGGCAAGGCGCGGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1332
 QY 1329 GTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
 DB 1333 GTTGAAGACACCGCTTCCAAATGAGGAGGACAGCTGACGTTCAACGCAAGAGG 1392
 QY 1389 GAAATATTAAGTTTGCATGACGATGAGGATGATGATGATGATGATGATGATGATGATGAT 1448
 DB 1393 GAAATCAAGGCTGCATGACGCTGAGGAGGATGATGATGATGATGATGATGATGATGATGAT 1452
 QY 1449 GATTAAGGACAAACAGCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1508
 DB 1453 GATTAAGGCAAAAGAAAGCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
 QY 1509 CAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1568
 DB 1513 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
 QY 1569 CGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
 DB 1573 CGTTTGAATTTAAAGGAGGATTTGCTTGGTTTCAACGTTATGATGATGATGATGATGATGAT 1632
 QY 1629 GCAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1688
 DB 1633 GCGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1692
 QY 1689 AGATTTGTTCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1748
 DB 1693 AGATTTTACAAACGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1752
 QY 1749 TACACGCTGTTGGCGAAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1808
 DB 1753 TACACGCTGTTGGCGAAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1812
 QY 1809 TATTAACCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1868
 DB 1813 TACACGCTGTTGGCGAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1872
 QY 1869 GATATTTCCAAACAAAGATTAATTTTTCAGGAGTATGATGATGATGATGATGATGATGATGAT 1928
 DB 1873 AACATCAGGAAACAAACGCAACGATTTTTCAGGAGTATGATGATGATGATGATGATGATGATGAT 1932
 QY 1929 AATCCTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1988
 DB 1933 AATCCTTTAGAGAGCGGTTGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1992
 QY 1989 GATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2048
 DB 1993 GATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2052
 QY 2049 GCGGCTTTTCCGCAATGTTTCTTCAATGAGGAAATTTGACAGTATGATGATGATGATGATGAT 2108
 DB 2053 GCGGCTTTTCCGCAATGTTTCTTCAATGAGGAAATTTGACAGTATGATGATGATGATGATGAT 2112
 QY 2109 AATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2168
 DB 2113 CAAGCAGTTTTTGGTGTGACCGCATCAAGGATGATGATGATGATGATGATGATGATGATGATGAT 2172
 QY 2169 ACGATTTAACGATTTGTCAAAAAGTATTTAACGATTTAACGATTTAACGATTTAACGATTTAAC 2228
 DB 2173 ACGGCTGACAAATTTGTGTCAAAAAGTATTTAACGATTTAACGATTTAACGATTTAACGATTTAAC 2232
 QY 2229 CCAAAAACCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2288
 DB 2233 ACTTAAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
 QY 2289 GATTTAGCAAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2348
 DB 2293 GAGCTTGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2352
 QY 2349 AACATGCAACCCAAATAGGCAATTTG----- 2377
 DB 2353 CACAAAGCACCACAAACGCAACCTTAGCTGTGGGCAATGCCAAGCAATTTAAT 2412
 QY 2378 ----- 2377
 DB 2413 CAAGCAGATTAAGGCAACATGCTTGGGCAATGCTTCAATTTATCTTAAACAC 2472
 QY 2378 -----ACTTTCCGACATTTGACATGCAACGTTGATTAAT 2411
 DB 2473 AACCGCGCAAAAAGGCAAGTCTGACGCTTTCCGACAGCTTAAGCAAAAGTAAACCAT 2532
 QY 2412 GCAAACTTGAACGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2471
 DB 2533 TCGGCACTCAAGGCAATGCTTCCCTAGCCGATTAAGGCAATTTCCATTTTGAACAC 2592
 QY 2472 CATTTTGCACCAAAATTCAGGAGACAAAGGCAACAGTACCTTGAAGAAATGCGACT 2531
 DB 2593 CGCTTTACCGACAACTCAGCGCGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
 QY 2532 TGGACATGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2591
 DB 2653 TGGACCTGCTCGCTAGGCAAGGATTTAACTTTGAACGACCAACCATTTACA 2712
 QY 2592 TTAATTTAGCTTAT-----TCAGTATGCTCAAAACATTAAG 2627

Db 2713 CTCGATTCGCGCTATTCGACGATGCTGAGCGCGCAACCGGCGAGTGTCTGACAGCG 2772
 Qy 2628 CCAAGTCGCGCGTTCAT-----TAGAGACGGAAACAAGCCAAACATCGGAGAAACAT 2678
 Db 2773 CCGCGCGCGCGTTCGCGCGCTTCCTTATTCGTAACCGCGCAACTTCGTAAGATCC 2832
 Qy 2679 CGTTTCACACATTCGACAGTAATAGTAATAGTGGGCAAGGCAATTCGTAATAC 2738
 Db 2833 CGTTTCACACGCTGACGTAAGGCAATTCGTAAGGCAATTCGTAATAC 2892
 Qy 2739 TCATCTTATTTGGCTATTAAGCGTAATTAATTAATTCATGTAAGCTGAGGCGCAT 2798
 Db 2893 TCGGAATCTTCGCTACCGACGACCAATTTAGAGCTGCGGAAGTTCGGAAGCACT 2952
 Qy 2799 TACATATATCTGTTCCGACACAGGCAAGACCGCAACCTTGAGCATTAACCTTG 2858
 Db 2953 TACACCTTGGCGGTCAACATACCGCAACGACCGCAAGCTGATCAATTAACGCTA 3012
 Qy 2859 GTTGAAGCAAGATATATCAACGTTATCAATGATGCTCAATTTACTTTAGAAATGAC 2918
 Db 3013 GTGGAAGGAAAGAAACAACACCGCTGCGAAACCTTAATTTACCTCGAAGAAACGA 3072
 Qy 2919 CAGTTGATGCAAGTGCATTAAGTTAATTAATTAATTAATTAATTAATTAATTAAT 2978
 Db 3073 CAGTGTATGCGCGCGGTGCGTTACCACTACCGCAAGACGCGAGCTTCGCGCTG 3132
 Qy 2979 CATACCCCAATTAAGACGAGAAATGCAACATGATTTAGTAAGACGAGACGAGAA 3038
 Db 3133 CATATCCGCTCAAGAAACAAAGCTTTCCGCAAACTCGGCAAGCGAAGAACCAAAA 3192
 Qy 3039 CGAATTTAGAACCA---ACAAGTTGAACGCACTGCTAACAACAACAGTGAACCA 3095
 Db 3193 CAGCGGAAAGAAAGCAACGCGCAACGCTTACGCGCTGATTCGCGCGCGGAGGCGC 3252
 Qy 3096 AAGTCGCGTCAAGAAAGCAAGCAAGAGAGGCTTTCTGATACCTGCTGATCAAGC 3155
 Db 3253 GCCGAAAGACAGAAAGGTTGCCGAACCGCGCGCGCGAGCGCGGAAAAAGTCCG 3312
 Qy 3156 CTGTAAAGCATTAAGAGCCAAACAGGCTAAGCTGACCTGTAACCAAAAAAGTAAG 3215
 Db 3313 ATTATCGAGCGGAGAAAGAAAGAAAGCGGTGACGGGAGT----- 3354
 Qy 3216 GCAAAAAAAGTACGCGTCAAAAAGAGAGTGTTCGATCCCTGCTGATCA 3275
 Db 3355 -----AAGACAGCGCTTGTGCGAAACAGCGCAAGCGGAAACCGCGCTACCA 3406
 Qy 3276 AGCTGTGCTATTAAGAACCGCACTTGAGGTTATGATCCGACAGCATCGAAAAA 3335
 Db 3407 CCGCTTCCCGCGCGCGCGCGCGCGCGCGCGGAGT-----TGCCTGCAACCGCACCGC 3459
 Qy 3336 GATGCTCTAGCTCAAGAAAGAGCGAAAAACAACGCAACAAAGAGCTTGACGCGCT 3395
 Db 3463 CAAACCGAACTCAACCCCAACG-----CAGCGCACTGATCAAGCGCT 3504
 Qy 3396 TATTCAAATAGTGCCTTATCAATTAATTCGCAAGTAATTAATTAATTAATTAAT 3455
 Db 3505 TATGCAATAGCGGTTGATGAAATTTCCGCAAGCTCAAGAGCTTTGCGCGTACAG 3564
 Qy 3456 GATGAATTAATGATGCTTTTGTGATCAAGCAATTCGCTGCTGCAAAATATGCA 3515
 Db 3565 GACGAATGGAACCGGTGTGCTGCGAAGACCGCGCAACCGCTTTGCAAGCGGAGTC 3624
 Qy 3516 CAGGATTAAGAGCGTATGATTCGATGCGTTCCGTTATATAGCAGCAAGAAACCAAC 3575
 Db 3625 CCGGACACCAAAACCTACCGTTGCAAGTTTCGCGCTACCGCAAC---AAACCGAC 3681
 Qy 3576 TTAAGTCAATTTGCGGTCAAAAAGCTTAAGCAATTAAGCAAGTTGCGGCAATTTTCG 3635
 Db 3682 CTGCGCAATCGGTATGCAAAAAACCTCGGCAAGCGGCGCGCTGCTGTTTCG 3741
 Qy 3636 CATAGCGCTTCAAGATTAATCTTTGATGAACAGGTTAAATTAATCAACGCAATTAAGCATG 3695
 Db 3742 CACAAACGAGACGAAAAACCTTCGAGAGCGCATTCGCAACTTCGCGACGGCTTGCAC 3801

Qy 3696 ATGTGCGGTTTCCCAATATCAATGCGGCGATTTACATTTGTGTAAAGTGGAAACG 3755
 Db 3802 GCGCGCGTTTCGGGCAATACGCGCATCGGCAAGGTTGACATCGGCATACACGCGCGCG 3861
 Qy 3756 GGAATGCTGCGATTAATTTGCTGAAAGAAACAAGCCAAAAATTCATCGAAAAAGATA 3815
 Db 3862 GGTTTTACAGCGCGAGCTTTCGAGAGGCGATCGAGGCAAAATCCGCGCGCGCTG 3921
 Qy 3816 AATTAAGCGTGAATGCAATTAACGTTACGTTGAGGCAATTTGAGCATTCAGCTTAT 3875
 Db 3922 CATTAAGCAATTAAGCAATTAACGTTGAGGCGGTTTTCGCGGATTCGCAATCGATC 3981
 Qy 3876 TTGAGTATATGCTATTTTATTAAGAGTGAATTAATCAATCTGAGAGTGAAGTG 3935
 Db 3982 ATCGCGCAACGCGCTATTTGCTCCAAAAAGCGATTAACGCTACGAAAGCTCAATAC 4041
 Qy 3936 AAACGCTGAGCTTCGATTAATTCGCTATTAATGCTGCGATTCGATTAATTAATTT 3995
 Db 4042 GCGACCGCGCTTTCGCTTCAACGCGTACCGCGCGCGGCAATTAAGGCAATTAATTC 4101
 Qy 3996 ACTCGGACAGTAATTAATGAGGTTAGGCTTATTTCTGCAATTAATGTTGATTTCA 4055
 Db 4102 AAACGCGCGCAACATTTCAATCAAGCTTATTAAGGCTTATTAAGGCTTATTAAGG 4161
 Qy 4056 AAGCTAACGTACAAACAAGGTAATTCACGCTGTTGCAACAACCATTTGACGTTAT 4115
 Db 4162 TCGGCAAAAGTCCGCAACGCGCTCAATACGCGTATTCGATTTGCGCAAAAC 4221
 Qy 4116 TCGCAAAAGAGTGGATTAAGGCAAGAAATTTTCAATTCGAAATTTCCGCTTTATC 4175
 Db 4222 CCGATGCGAATGCGGCGTAAACCGCAATCAAGGTTTACGCTGCTCCGCAAGCT 4281
 Qy 4176 TCAAAATCTAAGTTCAACATCGGCAACACGAAATAGTGGGCTGAAATGGGCTAT 4235
 Db 4282 GCGCGCGCAAGGCGCGCACTTGAAGCGCAACAGCGCGGCGCTCAATTAAGCTAC 4341
 Qy 4236 CGTTGTTAA 4244
 Db 4342 CCGCTGTA 4350
 RESULT 10
 ID AA12252 standard; DNA; 4350 bp.
 AC AA12252;
 DT 08-OCT-1999 (first entry)
 DE Neisseria meningitidis strain A complete ORF1 sequence.
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.
 OS Neisseria meningitidis.
 PN W09924578-A2.
 PD 20-MAY-1999.
 PF 09-OCT-1998; 98MO-IB001665.
 PR 06-NOV-1997; 97GB-03023516-
 PR 14-NOV-1997; 97GB-00024190-
 PR 18-NOV-1997; 97GB-00024386-
 PR 27-NOV-1997; 97GB-00025158-
 PR 10-DEC-1997; 97GB-00026147-
 PR 14-JAN-1998; 98GB-00000759-
 PR 01-SEP-1998; 98GB-00019016-
 PA (CHIR-) CHIRON SPA.
 XX XX

PI Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
 XX WPI: 1999-327407/27.
 DR P-PSDB; AAY38824.
 XX
 PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 PT diagnosis, treatment and prevention of infection.
 XX
 PS Claim 9; Page 364-365; 524pp; English.
 CC Nucleotide sequences AAL21972-212358 represent open reading frames (ORFs)
 CC of *Neisseria meningitidis* and *N. gonorrhoeae* which encode antigenic
 CC proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
 CC their nucleic acids and antibodies are used for diagnosis, prevention (as
 CC vaccines) or treatment of *Neisseria* infections, such as meningitis,
 CC septicemia and gonorrhea. Both organisms are closely related. Fragments
 CC of the nucleic acids are useful as hybridisation probes and antisense
 CC reagents
 XX
 SQ Sequence 4350 BP; 1225 A; 1161 C; 1060 G; 842 T; 0 U; 62 Other;
 Query Match 31.2%; Score 1348; DB 2; Length 4350;
 Best Local Similarity 59.0%; Pred. No. 0;
 Matches 2551; Conservative 0; Mismatches 1565; Indels 210; Gaps 9;

QY 801 TTTATTTATGATGCTGAAAACAAAATGCTTAATTAATGAGATTAATGCGAGGCAAC 860
 DB |||||
 QY 793 TTTATTTATGACAAAACAAACAAATTAATGCTGCTCAACGAGTTTAAACAAACGGCTAC 852
 DB |||||
 QY 861 CTTTGAAGCAAAAGAAATGGGTTTAATGTTGCGAATCTTATTT---TATGAA 917
 DB |||||
 QY 853 CTTTATCCCGACAGAAAACGGTTTCCAGCTATAGGAAAGATGGTTCTTACGATGAC 912
 DB |||||
 QY 918 ATTTCGAAAGAGATTAATGATACATCACTTAACCCGAGCTGTATGAGAGTGCACA 977
 DB |||||
 QY 913 ATTTCGAAAGAGATTAATGATACATCACTTAACCCGAGCTGTATGAGAGTGCACA 972
 DB |||||
 QY 978 ATTAGTGAAGATTAATGATGCTGAGGGTCTTCACTGAGAAATGAGATTAATGAGAA 1037
 DB |||||
 QY 973 TTATGATCCAAACAAACGGTACGGTACGGTAAAGAAACAAAGAAAGTNTCCAT 1032
 DB |||||
 QY 1038 ATTAAATTTAGCTTGAAGAAATATGAGTTT-----ACCTTTGAAAGAGAGTAA 1088
 DB |||||
 QY 1033 CAAAGCTTAAGTACAGAGAGTCCAGCTTTTACGAAATCTTTAATGAAACTGATATA 1092
 DB |||||
 QY 1085 GTTCATATCTCATATGACGAGACCTTAATTTATTTCCAGCTTTAAACAAATGAGAA 1148
 DB |||||
 QY 1093 GAACCAATTTACGGGGAGGGGGTGTATATGATACCGTCCAGGTTAAACAAAGGTGA 1152
 DB |||||
 QY 1149 AGCTTAATTTATGATGATCAAAACAAAGATCATTAATCTTGCATCTGACATTAACCA 1208
 DB |||||
 QY 1153 AACCTTCTTTTATGATATGACGCAACGGCAACCTATCTTATCAACAAATCAACCA 1212
 DB |||||
 QY 1209 GGGGGGGGTGCTTATTTTGAAGGTAAATTTAAGTATCTTCAATTTCCAAATTTAAACAA 1268
 DB |||||
 QY 1213 GGGGGGGGTGCTTATTTTGAAGGTAAATTTAAGTATCTTCAATTTCCAAATTTAAACAA 1272
 DB |||||
 QY 1269 TGGCAAGAGCTGACATACATGTAAGTAAATGACACCTTACTTGGAAAGTAAATGAC 1328
 DB |||||
 QY 1273 TGGCAAGAGCTGACATACATGTAAGTAAATGACACCTTACTTGGAAAGTAAATGAC 1332
 DB |||||
 QY 1329 GTGAAACATGATGATCTTAAATTTGTAAGAAACATGACCTTCAAGCCAAAGG 1388
 DB |||||
 QY 1333 GTGGCAACGACCGCGCTGCAAAATTCGGCAAGGACCTGCAAGTTCAAAGCAAGG 1392
 DB |||||
 QY 1389 GAAATTAAGTGTGATGACGGCTGAGAGGAGTAAAGTATTTGGAGACAGAGGAGAC 1448
 DB |||||
 QY 1393 GAAACCAAGCTGATGACGGCTGAGAGGAGTAAAGTATTTGGAGACAGAGGAGAC 1452
 DB |||||
 QY 1449 GATCAAGCAACAAACAGCTTTAGTAAATGCTTGGGTTAGGCGAGAGGAGCTGT 1538
 DB |||||
 QY 1453 GATTAAGCAACAAACAGCTTTAGTAAATGCTTGGGTTAGGCGAGAGGAGCTGT 1512
 DB |||||
 QY 1509 CAATTAACGATGATTAACATTTGATACGATTAATTTATTTGGCTTCTGCTGCT 1568
 DB |||||
 QY 1513 CAATTAACGATGATTAACATTTGATACGATTAATTTATTTGGCTTCTGCTGCT 1572
 DB |||||
 QY 1569 CGCTTAAGTCTTAAGGAGGCTCATTAACCTTAAAGTATCAACAAATTAAGGAGG 1628
 DB |||||
 QY 1573 CGTTGAGATTAAGGAGGCTCATTAACCTTAAAGTATCAACAAATTAAGGAGG 1632
 DB |||||
 QY 1629 GCAATGATTTGAACCAATTAACATCTMACCGCTATATGATCAATTAATCTGAGAA 1688
 DB |||||
 QY 1633 GCAATGATTTGAACCAATTAACATCTMACCGCTATATGATCAATTAATCTGAGAA 1692
 DB |||||
 QY 1689 AGCATTTGTTCTAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1748
 DB |||||
 QY 1693 AGTATTAACCAACAGAGTGTAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1752
 DB |||||
 QY 1749 TACAAAGTGTGTTGGCGAAACAGATAAATAACAAAGAGGCTTAAACCTTAT 1808
 DB |||||
 QY 1753 TACAAAGTGTGTTGGCGAAACAGATAAATAACAAAGAGGCTTAAACCTTAT 1812
 DB |||||
 QY 1809 TATTAACCAACAGAGATGATCTTGTACTTCAAGTGTGATCAATTTAAAGGC 1868
 DB |||||
 QY 1813 TACCAAGCTGCGGCAAGAGAGCGCAACCGGCTGTTCCGGGGAACAAATTTAAAGGC 1872
 DB |||||
 QY 1869 GATTTATCCCAACAAAGATTAATTAATTTTACAGGCTGATGACAGCAACCGGCTAC 1928
 DB |||||

Db 1873 AACATAGCGAACAAGCGAAGCTTTTTCAGCGGAGACCGACACCGCAGCTCAG 1932
Qy 1929 AATCATTTAATAAGCTGGTGTAGAAATGAGAGATACACAGCGGAAATTTGTGG 1988
Db 1933 AATCATTTAGAAAGCGGTGTGTCAAAAATGCAAGGTATCCCAACAGGAAATTTGTGG 1992
Qy 1989 GATCAGATTGATACACCGTACATTAAAGTGAAACTTCAATTAAAGCGGAAAT 2048
Db 1993 GACAGAGCTGATACACCGGATTTAAAGGAAATTTCAATTCAAGCGGAG 2052
Qy 2049 GCGGTGTTTTCGCAATGTTTCTCAATTGAGGAAATTTGACAGTACGCAATATGCA 2108
Db 2053 GCGGTATTTCCCGAATGTTTGCAGAAAGTGAAGCGATTGATTTGAGCAATACGCC 2112
Qy 2109 AATGCCAATTTGGTGTGTCGCAATACAAATAATCAATTTGACGCGTTCAATTGG 2168
Db 2113 CAAGCAGTTTGGTGTGTCGCAATACAAAGCAATCAATCTGTATCACGTTGCGAATGG 2172
Qy 2169 ACGGATTACGACTGTGTCAAAAATGCAATTTAAACGATACAAAGTATTAATCTATA 2228
Db 2173 ACGGTTGACAAATTTGTGTGCAAAAATCAATTAACGATTAAGTATTTCTCATTTG 2232
Qy 2229 CCAAAAACACAAATGCAATGCTCTATTAATTTAACTGATTAAGCAAGCGGCAATGTAAA 2288
Db 2233 ACTAAGACGACATTAAGGCGCAATGTAGACCTTACGATTAAGTATTTGATTTG 2232
Qy 2289 GGTTTAGCAAACTTAATGCAATGCACTTTAAACAAATACAGCCATTTACTTAAGC 2348
Db 2293 GGGCTGTGCACTTAATAGGCAATCTTAAGTGAATGAGCATACAGTTATACGTCAGC 2352
Qy 2349 AACAATGCCACCCAAATAGGCAATTTGCG----- 2377
Db 2353 CACAAAGCGACCCAAAGCGAAGCTTAAGCTGTGAGCAATGCCAAGCAACTTAAT 2412
Qy 2378 ----- 2377
Db 2413 CAAGCAGCTTAAGCGGCAAGCATGTGAGTTCGAGCAATTAATCTAAGCAAC 2472
Qy 2378 ----- 2411
Db 2473 AAGCGCGACAAAAGCGAGCTGTAGCGCTTTCCACAAAGCTTAAGGCAAGTAAAGCAAT 2532
Qy 2412 GCAAACTTGAAGGATATGTGATTAAGATTCAGTCAATTTCTTTAAAGAAAGC 2471
Db 2533 TCCGCACTCAAGCGCATTTCTCCCTAGCGATTAAGCATTAATTTGAAAGAGC 2592
Qy 2472 CATTTTGCACCAAAATTTCAAGGAGCAAAAGCGCAACAGTGAAGTTGAAATGCGACT 2531
Db 2593 GCGTTAACCGACAACTCAGGCGGACGAGCAAGGAAAGCATTAACATTAAGACAGGAA 2652
Qy 2532 TGACCAATGCGTGAAGATCTAATATTGCAAGATTTAAAGTTAATACAGTACAGC 2591
Db 2653 TGAAGCGTCCGTGAGGACGGAATTAAGCAATTTAAACCTTACAGCGCAACATTAACA 2712
Qy 2592 TTAATTCAGCTTAT-----TCACTAGCTCAAAAGCAATAGC 2627
Db 2713 CTCAATTTCCGCTTATGCGACAGATGCTGCAAGCGGCAACGCGGAGGTGACAGAGC 2772
Qy 2628 CCAAGTCCGCTTCAAT-----TAGAGCGGAAACAAAGCGCAACATGCGAGAAACAT 2678
Db 2773 CCGGCGCGCGCTTCCGCGCTTCCCTAATATCCGTTACACCGCAACTTCGATGAGATCC 2832
Qy 2679 CGTTTCAACATTTGACAGTAAATGCTAAATTTAGTGGGCAAGGCAATTTCACTTAAT 2738
Db 2833 CGTTTCAACAGCTGACAGTAAAGCGCAATTTGAACNGTCAAGAAACATTTCCGCTTAAG 2892
Qy 2739 TCAATCTAATTTGGCTAATAAAGCATATAATTAATTAATCAATGACGTGAGGCGCAT 2798
Db 2893 TCGGAACCTTTGGGTACGAGAGCGCAATTTGAAGCTGCGGAAAGTTCCGAGAGAACT 2952
Qy 2799 TACATTAATTTGTTGCAACAGAGGCAAGAAAGCGGAAACCTTTAGCAATTAATTTG 2858

Db 2953 TACACTTTGGCGGTCAACAAATACGAGCAAGCAAGCGGTAAGCTGATCAATGAGGTA 3012
Qy 2859 GTTGAAGCAAGATTAATGACCGTTATCAGATTAAGCTCAATTAATTAAGAAATGAC 2918
Db 3013 GTGGAAGGAAAGACAAACAAACCGGTGTCCGAAATCTTAATTTACCTTCAAAACGAA 3072
Qy 2919 CACGTTATGACAGGTGATTAAGTTAATTAATTAAGTGAAGAAATGATGCGAATTCGCTTG 2978
Db 3073 CACGTGATGTCGCGCGCGGTGTGTTACCACTCATTCGCAAAAGAGCGGATTCGCGCTG 3132
Qy 2979 CATTAACCAATTAAGAGCAAGAAATTCACATGATTTATGTAAGAGACGACAGAGAA 3038
Db 3133 CATATCCGCTCAAGAAACAAAGAGCTTTCCGACAAACTCGGCAAGGC----- 3179
Qy 3039 CCAACATTTAGAGCCAAACAAAGTGAACGATCTTAAACCAACAAAGGTGAGCCAAA 3098
Db 3180 -----AGAGCGCAAAACAAAGCGGAAAGCAACGCGCAAAACCTTTGAGCGCTG 3221
Qy 3099 GTGCGGTCAAGAGACAGAGAGCAGCGTTTCTGATACCTGTGCTGATCAAAAGCTG 3158
Db 3232 ATTTGCGCGCGGCGCGATGC---CGCGAAAGACAGAAAGCTTTGCGAATCCGCGCGG 3288
Qy 3159 TTAACGCAATTAGAACCCAAACAAAGCTGAATGCTGTAAGAAACAAAGAAATAGGCA 3218
Db 3289 CAGCAGCGCGGGAATAATGTCCGCAATTATCAGCGCGAGAGAAATAACCGGTGACG 3348
Qy 3219 AAAACAAATAAAGTGGCGTCCAAATAAGCAGTGTTCGTGATCCCTGCTGATCAAAAGC 3278
Db 3349 GCGGATTAAGACAGCGCTTTGCGAAACAGCGAAAGCGGAAACCGCGGTACCAAC 3408
Qy 3279 CTGTTGCAATTAAGACCGCACTTGAAGTTATGATGCCCCACAGCAATGCAAAAGAT 3338
Db 3409 GCGTTCCCGCGCGCGCGCGCGCGCGCGCGGATTTTCCGCAACCGCAACCGCAACGCA 3458
Qy 3339 CGCTAGCTTAAGAGAGAGCGGAAACAAAGCAACAAAGAAAGCTTGATACCGCTTAT 3398
Db 3469 CCTCAACCCCA-----CGCAGCGCGACCTGATTAAGCTTAT 3507
Qy 3399 TCAATATGCGTTATCAGATTAATCTGCAACAGTAAATGATAGCTTGTGTAAGAT 3458
Db 3508 GCGAATAGCGGTTTGAATTAATTTCCGCAAGCTCTCAAGAGGTTTTCGCGTAAGAGAC 3567
Qy 3459 GAATTAAGTGTCTTTTGTAGATCAAGCAACATCTGCGGTGTGACAAATATGCAAGC 3518
Db 3568 GAATTAAGCGCGGTTTTCGCAAGACCGCGCAAGCGGATTTGACACACGCAATCCG 3627
Qy 3519 GATTAAGAGCGCTATGATTTGATGCTGCTTTCGCTTATCAGAGCAAGAAAGAACTTA 3578
Db 3628 NAACCAAAACATACGCTTCCGCAATTTCCGCGCTTACCGCAAC---AAACCGACCTG 3684
Qy 3579 CGTCAATTTGGGCGTGGCAAAAGCTTATGATGACGAATTTGGGCGAGTTTCTGCAAT 3638
Db 3685 CCGCAATGTGATGCAAGAAACCTTGCAGCGGCGGCGGTGCGATCTGTTTTCGAC 3744
Qy 3639 AAGCGTTGATTAATCTTTGATGAACAGGTTAAATCAAGCGACATTAAGCATGATG 3698
Db 3745 AACCGGACGAAACAACTTTGACAGCGGCAATCGCAATCGGCAACGCGCTTGCCACGCG 3804
Qy 3699 TCGGTTTTCGCAATTAATGATGAGGCGATTTCAATTTGCTGTAACGTTGGAAGGGA 3758
Db 3805 GCGGTTTTCGCGCAATTAAGGCAATGCGAGGTTGACATGAGATGACAGCGGCGGCT 3864
Qy 3759 ATCAATGCGATTAATTTGCTGAGAACAAAGCGCAAAATCTCAATGAAAGGATTAAT 3818
Db 3865 TTTAAGAGCGCAATCTTATCAAGCGATGAGAGGAAATTCGCGCGCGCGCTGAT 3924
Qy 3819 TATGCGGTGAATCAAGTATCAATGCTTTTATGAGCAATTTGAGCAATTAAGCTTATTT 3878
Db 3925 TACGCAATTTGAGGACAGATACGCGCGGTTTTCGCGGATTTGCGATCGAATCGTATC 3984
Qy 3879 GAGGTTATGCTATTTTATGAGAGGAAATTAATCAACTGAGGAAAGGAGAGGAA 3938
Db 3985 GCGGCAAGCGCTATTTGCTCAAAAGCGGATTTACCGCTAGGAAAGGTAATATGCGC 4044

QY 3939 AGGCTAGCTTSCATTTATGCTATTAAGTGGGATTCAGTGAATTAACAATTACT 3998
 DB 4045 ACCCCCGCTTTCGGTTCAACGATACCGGCGGATTAAGGAGATTAATTCATTCAAA 4104
 QY 3999 CCGACAGATTAATTCAGCGTTAAGCTTATTTCTTCGTCATTAATGATGTTTCAAC 4058
 DB 4105 CCGGCGACAGATTAATTCAGCGTTAAGCTTATTTCTTCGTCATTAATGATGTTTCAAC 4164
 QY 4059 GCTAACGTAACAACAGCGTAAATTCAGCGTGTTCGACACACCTTTGACGTTATGG 4118
 DB 4165 GCGAAAGTCGAAACAGCGCTCAATACCGGATTTGCTCAGATTTCCGCAAAACCCGC 4224
 QY 4119 CAAAAGAGTGGGATTAAGGCAAGAAATTTTACATTTCCAAATTCGCTTTATCTCA 4178
 DB 4225 AGTGGGATGAGGCGGTAAGCGCGAATCAAGGTTTACGCTGTCTCCATCCAGCGTCC 4284
 QY 4179 AATCTCAAGTTCACACTCGGCAACAGCAAAATGAGGCGCTGAATTTGAGCTATGCT 4238
 DB 4285 GCGCGCAAGGNCGCAACTGGAAGCCCAACAGAGCGCGGCTATCAATTTAGCTACCGC 4344
 QY 4239 TGGTAA 4244
 DB 4345 TGGTAA 4350

RESULT 11

AAZ12253

ID AAZ12253 standard; DNA; 4407 BP.

XX AAZ12253;

XX 08-OCT-1999 (first entry)

DE Neisseria gonorrhoeae complete ORF1 sequence.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

XX treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.

OS Neisseria gonorrhoeae.

XX MO3924578-A2.

XX 20-MAY-1999.

PE 09-OCT-1998; 98MO-IB001665.

PR 06-NOV-1997; 97GB-00023516.

PR 14-NOV-1997; 97GB-00024190.

PR 18-NOV-1997; 97GB-00024386.

PR 27-NOV-1997; 97GB-00025158.

PR 10-DEC-1997; 97GB-00026147.

PR 14-JAN-1998; 98GB-00000759.

PR 01-SEP-1998; 98GB-00019016.

XX (CHIR-) CHIRON SPA.

XX Masignani V, Rappuoli R, Pizzia M, Scarlato V, Grandi G;

XX WPI; 1999-327407/27.

XX P-PSDB; AAY38825.

XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for

XX diagnosis, treatment and prevention of infection.

XX Claim 9; Page 370-371; 524pp; English.

XX Nucleotide sequences AAZ11972-212358 represent open reading frames (ORFs)

XX of Neisseria meningitidis and N. gonorrhoeae which encode antigenic

XX proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,

XX their nucleic acids and antibodies are used for diagnosis, prevention (as

XX vaccines) or treatment of Neisseria infections, such as meningitis,

XX septicaemia and gonorrhea. Both organisms are closely related. Fragments

CC of the nucleic acids are useful as hybridisation probes and antisense
 CC reagents
 XX
 SQ Sequence 4407 BP; 1273 A; 1169 C; 1094 G; 871 T; 0 U; 0 Other;

Query Match 30.0%; Score 1297.8; DB 2; Length 4407;
 Best Local Similarity 59.1%; Pred. No. 6.4e-303;
 Matches 2571; Conservative 0; Mismatches 1567; Indels 213; Gaps 12;

QY 86 TTTTAAACCGCTTGCTTTCATTTAGGATGATGCGAAGTGGGTGCTGACCTTA 145
 DB 78 TTACTTACGATATGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 137
 QY 146 TTTTGGATGATATACCAATATATGATGATTTGCGGAAATTAAGGAGATTCACAGT 205
 DB 138 TTTCGCGATCACTACCAATATATGATGATTTGCGGAAATTAAGGAGATTTGAGT 197
 QY 206 TGGGCTCAAAATATTAAGTTTAAACAAAGGCAATTAAGTTGCGCATCATGATGAC 265
 DB 198 CCGGCGCAAAAGTATGAGGTTTAAACAAAGGCAATTAAGTTGCGCATCATGATGAC 257
 QY 266 AAAAGCCCGATGATGATTTTCTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 325
 DB 258 GAAAGCCCGATGATGATTTTCTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 317
 QY 326 AAATCAATATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 385
 DB 318 CGATCAATATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377
 QY 386 AGAGGGAACAAACCCGATCAACATGCTTTACTTATTAAGATTTGAAACAAATTA 445
 DB 378 GAGGGAACAAACCCGATCAACATGCTTTACTTATTAAGATTTGAAACAAATTA 437
 QY 446 CAAAAG---KTAATTAACATCTTAATGAGATTAACATTAATCAACATTA 502
 DB 438 TAAAGGAGGATCAACGCTATCTTAATGAGATTAACATTAATCAACATTA 497
 QY 503 ATTGCTTACAGAGGCTGCTCAATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTG 562
 DB 498 ATTGCTTACAGAGGCTGCTCAATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTG 557
 QY 563 AGATAGAACAAATATCCAGAGCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTG 622
 DB 558 TGATTTAAATTAATATCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
 QY 623 TGATCAAGCAAAAGC-----GACCAAGTGGCGGCTGATTA 664
 DB 618 TGATGAAGAGCAACCAATTAACCGGAAAGTTCAATATTAATGCAAGCATTA 677
 QY 665 TCTGACAGCTGGCAATACACATGAGCTGAGAGGATTAATGATTAATGATTA 724
 DB 678 GCTGTGCTGCTGCAATACCTTGTGCAAAATGATGAGGCTGAGCAAGTCACTA 737
 QY 725 AGGCGATGTTGTAAGGCGGAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784
 DB 738 TAGCGAAATTAATTAACATGAGCCATTAATGCTTTTTCACAAAGAGGCTATTTG 797
 QY 785 CAGTGTGCTGCTGATGATTAATGATGCTGAAACAAATTAATTAATTAATG 844
 DB 798 CAGTGTGCTGCTGATGATTAATGATGCTGAAACAAATTAATTAATTAATG 857
 QY 845 ATTACGGAAGGCAACCTTTTGAAGCAAAAGAAATGAGTTTCAATGTTGCGAATC 904
 DB 858 ATTACGGAAGGCAACCTTTTGAAGCAAAAGAAATGAGTTTCAATGTTGCGAATC 917
 QY 905 TTATTT---TGATGAATTTTGAAGAGATTAATGATGCTTACACCGAGCTG 961
 DB 918 TTGATTTATGATGAAGATTTTGTGAGATTAATGATGCTTACACCGAGCTG 977
 QY 962 TTAATGAGTGTACAAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1013
 DB 978 AATGGAATATCTTTTAAACAAATTAATGATGATGATGATGATGATGATGATGAT 1037

QY 1014 CAGAAATCGAATACCATGAGAAATTAATTAC-----TTAGCAATATGAG 1063
 DB 1038 TAAACCTATCTCTACCTTATGATTAATAACGACGCTCAATGTTAAATGTTTC 1097
 QY 1064 TTTACTTTGAAAGAGAAATTAAGTTCAATACCTTAAGATGACGACCTTAATTATTA 1123
 DB 1098 TTTATCCGAGACGAGAAAGAACTGTTTATATGCTGCA--GGTGGGCTCAAGATTA 1154
 QY 1124 TTCTCAGCTTTAAACATGAGAGAAACGCTATATTTTATGATCAAAAACAAGATCAT 1183
 DB 1155 TCGACCAACATGAAATATGAGAAATTTTCTTTATATGACAAAGAGAAAGGATTA 1214
 QY 1184 AATCTCGATCTGACATTAACAAGGCGGCTGCTTATTTTGAAGGTAATTTTAC 1243
 DB 1215 GATACCTACGACACATCAACCAAGCGCGGCTGCTTATTTGAGGTAATTTTAC 1274
 QY 1244 AGTATCTCCAAATTTCTAACAACTTGCGAAGAGCTGCGATCATATGATGAAATTA 1303
 DB 1275 GGTCTGCTTAAACAAACGAAAGTGGCAAGCGCGGCTCATATCATGATGCGAG 1334
 QY 1304 CACCGTACTTGGAAAGTAATGCGCTGGAACATGATGACTTTCTTAAATTTGTTAAAG 1363
 DB 1335 TACGCTTACTTGAAGAAAGTAACGCGCTGGCAACGACGCTGTCCAAATTCGCAAG 1394
 QY 1364 AACATTTGACGTTCAAGCCAAAGGAAATTAAGCTTCAATGAGCTGATGCTAA 1423
 DB 1395 CAGCTGCTGCTTCAAGCCAAAGGAAATTAAGCTTCAATGAGCTGATGCTAA 1454
 QY 1424 AGTATTTTGGACGACGACGACATCAAGGCAACAAAGCTTATGAAATTTG 1483
 DB 1455 AGTATCTTGAATCGACGACGACATCAAGGCAACAAAGCTTATGAAATTTG 1514
 QY 1484 CTGCTGAGCGGACGAGGACCTGTTCAATTAACGATTAACATTTGATCCGATTA 1543
 DB 1515 CTGCTGAGCGGACGAGGACCTGTTCAATTAACGATTAACATTTGATCCGATTA 1574
 QY 1544 ATTTATTTGGCTTGGTGGTGGCTTAATCTTAAGGCTATCTTAACCTTTAA 1603
 DB 1575 ACTATTTTGGCTTGGTGGTGGCTTAATCTTAAGGCTATCTTAACCTTTAA 1634
 QY 1604 ACATTTTCCAAATTAACGACGAGGCGATGATGATGACATTAATCAACCTCAAGCGC 1663
 DB 1635 CCGATTTCCAAATTAACGACGAGGCGATGATGATGACATTAATCAACCTCAAGCGC 1694
 QY 1664 TAATGCTATTTACTGAGAACGAAGCATTTGCTCAATTAAGAAATTAATTAATA 1723
 DB 1695 CACGCTTACATTTACGAGCAATTA-----GATATTTACTACAACGCGCAATTAACAA 1748
 QY 1724 ACTGATTTACAGAAAGAAATTTCTTAACGAGCTTGGTGGGAAACGATTAATAA 1783
 DB 1749 CTGGATGACAAAGAAATTTGCTTCAACGCTTGGTGGGCAAGAAAGATGCAACCA 1808
 QY 1784 ACACAAATGGCGATTAACCTTATTAACCAACGACGAGAGATGATCTTGGTACT 1843
 DB 1809 AACGACGCGCGCTCATCTGAATTAACCAACGAGAGAGCGATGCTTTACTCT 1868
 QY 1844 TTCAGTGTGTAACAATTTAAAGCGATTTTACCCAAACAAAGATTAATTTTAC 1903
 DB 1869 TTTCCGCGGAAACAAATTTTAACGCAATTAACGCAACAAAGGCAACTGTTTAC 1928
 QY 1904 CGGTAGACGACGACGCGCTTCAATCTTAATTAACGCTTGGTGGGAAATGAAAG 1963
 DB 1929 CGGACGACGACGACGCGCTTCAATCTTAATTAACGCTTGGTGGGAAATGAAAG 1988
 QY 1964 TATACCAAGGCGAAATTTGTGGGATCAACATTTGATCAACGCTTAAAGCTGA 2023
 DB 1989 TATCCCAAGGAGAAATGTTGTGGACAAACATTTGATCAACGCTTAAAGCGGA 2048
 QY 2024 AAACCTTCAATTTAAAGGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2083
 DB 2049 AAACCTTCAATTTAAAGGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2108
 QY 2084 AAATTTGACAGTCAATATATGCAATTTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2143

DB 2109 CAAATTTGACAGCTTGAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2168
 QY 2144 TACGATTTGACAGCTTGAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2203
 DB 2169 CAAATTTGACAGCTTGAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2228
 QY 2204 CAAATTTGACAGCTTGAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2263
 DB 2229 CAAATTTGACAGCTTGAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2288
 QY 2264 TCAATTTGACAGCTTGAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2323
 DB 2289 CAAATTTGACAGCTTGAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2348
 QY 2324 AAATCAGGCA----- 2335
 DB 2349 CGGACGACGACGATTAAGCTTACCGGCAACGCAACCCAAACGCAACCTGAGCTGG 2408
 QY 2336 ----- 2335
 DB 2409 GGGCAATGCCGACCAATTTATCAAGCCCAATTAACGCAACATGCGCTTGGCA 2468
 QY 2336 -----ATTATCAATTAAGCAACATGCGCAACCAATTAAGCAATTTGCACTTCCGA 2386
 DB 2469 CAATGCTTCAATTTATTAAGCAACATGCGCAACCAATTAAGCAATTTGCACTTCCGA 2528
 QY 2387 CAATTTCACTGCAACGCTGATTAATGCAACCTTGAACGTAATGCAATTTAAGCAATTC 2446
 DB 2529 CAACGCTAAGCAACGTAAGCACTTCCGCACTCAACGCAATGCTCCATGACCATTA 2588
 QY 2447 AGCTCAATTTCTTTAAACCAAGCAATTTTGGCAACCAATTTGAGGAGCAACAGCAC 2506
 DB 2589 GGCAGATTTCCATTTGAAACAGGCGCTTACCGGAAATTAAGCGCGGCAAGATTA 2648
 QY 2507 AACAGTACGTTGAAATGCACTTGAACATGCTGACATGCTGACATTAATTAAGCAATTT 2566
 DB 2649 GGCATTTACATTTAAGCAACAGCAATGACATGCTGCTGCGGCAAGATTTAAGCAATTT 2708
 QY 2567 AACGTAATTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2606
 DB 2709 AACCTTGAACGCACTTACACTCAATTCGCTTACGACATGCGGCGAGCGC 2768
 QY 2607 -----TGAGTACGCTCAACCAATTAAGCAATTT-----CGCGCTTATTAAGCAAGCA 2653
 DB 2769 GCAACCGCGAGTGGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2828
 QY 2854 AACGACGCAACATGCGGCAACATGCTTGAACATGATGATGATGATGATGATGATGATGAT 2713
 DB 2829 TACGCGGCAACCTTGGCAACATGCTTGAACATGATGATGATGATGATGATGATGATGAT 2888
 QY 2714 TGGGCAAGGCAATTTCACTTCACTTTATTTGGCTTAAAGCATTAATAA 2773
 DB 2889 CGGCTAGGGAACATTTCCGCTTATGCGGAACCTTCTGCTTACCGCAACGCAATTTGA 2948
 QY 2774 ATTTATCAATGACCTGAGGCGCAATTAATTAATTTGCTTGAACAGGCAACCAAC 2833
 DB 2949 GCGGCGGGAAGTTGCAAGGCACTTAACCTTGGCTTGAACATTAACGCAACCAAC 3008
 QY 2834 CCAACCTTGAACATTTACTTGGTGAAGCAATTAATTAACGCTTATCAATTA 2893
 DB 3009 CCAATGCTGAGCAATTTGAGGCTGATGAGGAAAGCAACACCGCTGTCCGAA 3068
 QY 2894 GCTCAATTTACTTGAAGCAATTTGAGGCTGATGAGGAAAGCAACACCGCTGTCCGAA 2953
 DB 3069 TCTTAATTTCACTTGAAGCAATTTGAGGCTGATGAGGAAAGCAACACCGCTGTCCGAA 3128
 QY 2954 GAAGATTTGAGGCAATTTGAGGCTGATGAGGAAAGCAACACCGCTGTCCGAA 3013
 DB 3129 CCGGAAAGCGGAGGCTTCCGCTGATTAATGCTGATTAAGCAACAGCTTTCCGAA 3188
 QY 3014 TTTATTAAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3073

Db 3189 ACTGCGAAGCGGAGGAAAAGAGGCCGCTTGAACGCGAACAAGCAACATTCGCGC 3248
 QY 3074 TAAATACGAAACAGGTGAGCCAAAAGTCCGTCAAGAAAGACGAGAGAGCGCTTCC 3133
 Db 3249 CAAACAAACGCGGAAAAAACAACGCGAAAACCTTGACGCTGATTTGGCGCGGCG 3308
 QY 3134 TGATACCTGCTGATCAAGCCCTGTAAACGCAATTGAAGCCCAACAGCTGAATGAC 3193
 Db 3309 CAATGCAACGAAAAGGCAAGAAAGTGTTCGGAACCGGCCCGGACGAGCGGAGAAA 3368
 QY 3194 TGTGAACAACAAAAAGTAAGCAAAAACAAAAAGTCCGTCAAAAAAGCAGATGTT 3253
 Db 3369 TGCCGCGATTATGACGCGGAGAGAGAAAAACGCGTCAAGCGATTAAGACACCG 3428
 QY 3254 TTCTGATCCCGCTGTATCAAAAGCTGTTGCAATTAAGACCGCACTTGAAGTTATGA 3313
 Db 3429 CTGGGGAACAGCGCGAGAGAAACCGCGCGCTACACACGCTTCCCGCGCGCG 3488
 QY 3314 TGCCCAACAGCAATCGAAGAAAAGATGCTTACCTCAAGAAAGCGAAGAAACAGCA 3373
 Db 3489 CCGGCGCGCGCGGATTTGCGGCAACCGAGCCCAACCGCAACCGCAACCGGACGCGC 3546
 QY 3374 ACAAAAAGACTGATCAGCGCTTATTCAAATATGTCGCTATCAAGATTTCTGCAACG 3433
 Db 3547 -----GACCTGATCAGCGCTTATGCAATGACGCTTGTGAGTGAATTTCCGCGACGCT 3599
 QY 3434 AAAGATAGCTTCTGCTCAAGTGAATTAAGTCTTCTTGTGATGCAAGCAACATC 3493
 Db 3600 CACAGCGTTTTCGCGTACAGCAATTAAGACCGCGCTTTCGGAAGACCGCGCA 3659
 QY 3494 TGCGTGTGACAAATATCGACAGATTAAGAAAGCGTATGATTCGATTCGCTG 3553
 Db 3660 CGCGGTTTGAACACCGGCAATCGGAGACACAAACAGTACCGTTGCAAGATTTCCGCGC 3719
 QY 3554 TTATCAGACGCAAAAGAACTTACGTAATTTGGGCGTCAAAAAGCTTACGTAATG 3613
 Db 3720 CTACCGCAAC---AAACGACCTGCGCAATCGATGACAAAAACCTGCGACGCG 3776
 QY 3614 AGCAATGCGGCGATTTCTGCAAGTACGCGCTTCAATTAATCCTTTGATGAACAGTTA 3673
 Db 3777 GCGGTCGATCTGTTTTCGCAACGAGACCGGAAACCTTTGACGAGCGATCG 3836
 QY 3674 AATCAGCGGATTAACGATGATGCGGTTTTCGCAATTAATCAATGCGGCGATTTA 3733
 Db 3837 CAATCGGACCGGCTTTCGCAAGGTCGCTTTCGCAATTAATCAATGCGGCGATTTA 3896
 QY 3734 ATTTGCTGAACGCGGAGAGGATGATGCGGATTAATGCTGAAGAAACAAACCG 3793
 Db 3897 CATCGGATCAGCGGAGCGGCTTGTAGTACGCGGCTTTCAGACGCGATCAGAG 3956
 QY 3794 AAAAATCATCGAAAAGGATTAATTAAGCGCTGAATGCAATTAATCAATGCGGTTA 3853
 Db 3957 CAAATCCGCGCGCGGCTGCTGATACGCAATTAAGGCAATTAAGGCGGCTTCCG 4016
 QY 3854 GCAATGCGGATCAGGCTTATTTGAGTAACTGCTATTTATTAAGTGAATTA 3913
 Db 4017 CGAATCGGATCAGGCAACGCAATCGCGGCAACGCGCTATTTGCTCAAAAAGCGATTA 4076
 QY 3914 TCAATCGAGAGAGTGAAGTGAAGAAAGCGCTTGAATTAATGCTGAATTAAGTGG 3973
 Db 4077 CCGATGAGAAAAGTCAATATGCGCACCCCGGCTTCAATTAACCGGCTCCGCGG 4136
 QY 3974 CATTGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4033
 Db 4137 CATTAAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4196
 QY 4034 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4093
 Db 4197 CCGTCTATACCGATGCGCTTCCGCAAGTCCGAAAGCGCTCAATTAATTAATTA 4256
 QY 4094 GCAACACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4153
 Db 4257 GCGGAGATTTTGGCAAAACCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 4316

QY 4154 TTTCGAATTTCCGCTTATATCAAAATCTCAAGTTTCAACTCGGCAACGCAAAA 4213
 Db 4317 TTTCAGCGCTTCTTCACAGCTGCGCGCGCAAGGCGGCAATTTGAAGCGAGACAG 4376
 QY 4214 TGTGGCGTGAATTTGAGCTATCGTTGTA 4244
 Db 4377 CCGGCGCATCAAAATTAAGCTTACCGCTGTA 4407
 RESULT 12
 AB238249
 ID AB238249 standard; DNA; 4404 BP.
 XX
 AC AB238249;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae nucleotide sequence SEQ ID 1087.
 XX
 DE Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN W0200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002MO-IB02069.
 XX
 PR 12-FEB-2001; 2001GB-0003424.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Fontana MR, Pizsa M, Masignani V, Monaci E;
 DR WPI; 2003-058415/05.
 DR P-PSDB; ABP77279.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 XX medicament for treating or preventing N. gonorrhoeae infection.
 XX
 PS Disclosure; Page 266-267; 815pp; English.
 XX
 CC The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records AB237706-AB242316 represent nucleic acid
 CC molecules of the invention
 XX
 SQ Sequence 4404 BP; 1271 A; 1170 C; 1094 G; 869 T; 0 U; 0 Other;
 Query Match 30.0%; Score 1294.8; DB 7; Length 4404;
 Best local Similarity 59.1%; Pred. No. 3,4e-302;
 Matches 2568; Conservative 0; Mismatches 1567; Indels 213; Gaps 12;
 QY 86 TTTTAAACGCTTGCATTTCAATGAGATGATGCGAAGCGTGGTCAACTTA 145
 Db 78 TTACTTACGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 137
 QY 146 TTTTGGATGATTAACCAATATATGATTTTGCAGAAATTAAGGAGATTCACAGT 205
 Db 138 TTTCGCAATCAACATACATATGCGACTTTGCGAAAAATTAAGCAAGTTGAGT 197
 QY 206 TGGGCTCAAAATTAATTAAGTTTATTAACAAAGGCGAATTAATGTTGCAATCATGAC 265
 Db 198 CCGGCGCAAAATTAATTAAGTTTATTAACAAAGGCGGAGTTGTTGCGCAATCATGAC 257
 QY 266 AAAGCGCGGATGATTAATTTTCTGTAGTCACTTAACGCGGCTGCTGTTGTA 325

Db 258 GAAGGCCGATGATGATTTTCTGTGTATCGGTAACGGCGGCGGATGCGCGG 317
 Qy 326 AATCAATATATTTGTGAGGTGGCACAATAGTATGATATCAAGTATTTTGTGTC 385
 Db 318 CGATCAATATATTTGTGAGGTGGCACAATAGTATGATATCAAGTATTTGTC 377
 Qy 386 AGAGGGAACCAACCGGATGATATGATTTTACTTATGATTTGTAACCAATATGTA 445
 Db 378 GAGGGAACCAATCCGATGATGATCGCTTTTCTTACCAATTTGTAACCAATATGTA 437
 Qy 446 CAAAAAG---ATAATTTACATCTTATGAGAGATTTACCAATTTCCAGATTAATTA 502
 Db 438 TAAAGCAGGACTAACCGGCACTCTTAATGCGGATTTACATATGCGCGCTTGACAA 497
 Qy 503 ATTGCTTACAGAGCGGCTCAATTTATGATCTTCAATATGATTTGCAATTTATTC 562
 Db 498 ATTGCTTACAGAGCGGCTCAATTTATGATCTTCAATATGATTTGCAATTTATTC 557
 Qy 563 AGATTAACCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 622
 Db 558 TGATTTAATTAATATCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 617
 Qy 623 TGATCAAGCAAAAGC-----GACCAAGTTGCGGCTGATATCAATTA 664
 Db 618 TGATTAAGCAAAAGC-----GACCAAGTTGCGGCTGATATCAATTA 677
 Qy 665 TCTGACAGCTGGCAATACACATAGGCTGAGAGCAATATGATTTGCTATG 724
 Db 678 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
 Qy 725 AGGCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784
 Db 738 TAGGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 797
 Qy 785 CAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 844
 Db 798 CAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
 Qy 845 ATTAAGGGAAGCAACCTTTGAGAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 904
 Db 858 ATTGCAAGAGCAACCTTTGAGAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 917
 Qy 905 TTAATTT---TGATGAATTTTCAAGAGATTTATCAATCACTTACACCCAGCTG 961
 Db 918 TTGCTTATGATGAATTTTCTGCTGAGATACCATTTGATTTCTAGAACACATCA 977
 Qy 962 TAAAGAGTGTACCAATTAAGTGAATGATATGATGATGATGATGATGATGATGAT 1013
 Db 978 AATGGAATTAATTTTCAAGCAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1037
 Qy 1014 CAGGAATCAGGAATACATCAAGAAATTAATTAAG-----TTAGCAATTAAGAG 1063
 Db 1038 TAAACATTAATTTCTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1097
 Qy 1064 TTAATCTTGAAGAGAGATTAAGTATTAATTAATTAATTAATTAATTAATTAATTA 1123
 Db 1098 TTTATCCGAGACAG 1154
 Qy 1124 TTTCCAGCTTTAAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
 Db 1155 TCGACCAAGCTGAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1214
 Qy 1184 AATCTTGCATCTGACATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1243
 Db 1215 GATCTTACAGCAATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1274
 Qy 1244 AGTATCTCAAAATTTTAAACCAAACTTGGCAAGAGCTGCTCAATGATGAATTAAG 1303
 Db 1275 GGTCTGCTTAAAG 1334
 Qy 1304 CACCGTTACTTGAAG 1363
 Db 1335 TACGTTACTTGAAG 1394

Qy 1364 AACATTCAGCTTCAAGCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1423
 Db 1395 CACCGTTACTTGAAG 1454
 Qy 1424 AGTATTCAG 1483
 Db 1455 AGTATTCAG 1514
 Qy 1484 CTGCTTGAAG 1543
 Db 1515 CTGCTTGAAG 1574
 Qy 1544 ATTATTTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603
 Db 1575 ACTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1634
 Qy 1604 AGTATTCAG 1663
 Db 1635 CCGATTCAG 1694
 Qy 1664 TAAATGCTATTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1723
 Db 1695 CACCGTTACTTCAAG 1748
 Qy 1724 ACTTGAATTAAG 1783
 Db 1749 CTGCTTGAAG 1808
 Qy 1784 ACAATGAG 1843
 Db 1809 AAGCAAG 1868
 Qy 1844 TTCAGTGTATCAAAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1903
 Db 1869 TTCAGTGTATCAAAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1928
 Qy 1904 CCGTGAAG 1963
 Db 1929 CCGTGAAG 1988
 Qy 1964 TATACCAAG 2023
 Db 1989 TATACCAAG 2048
 Qy 2024 AATCTTCAAAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2083
 Db 2049 AATCTTCAAAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2108
 Qy 2084 AATCTTCAAAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2143
 Db 2109 AATCTTCAAAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2168
 Qy 2144 TATACCAAG 2203
 Db 2169 TATACCAAG 2228
 Qy 2204 CAGATCAAAAG 2263
 Db 2229 CAGATCAAAAG 2288
 Qy 2264 TATACCAAG 2323
 Db 2289 CAGATCAAAAG 2348
 Qy 2324 AATCTCAAGCA----- 2335
 Db 2349 CAG 2408
 Qy 2336 ----- 2335
 Db 2409 GGGCAATGCCCAAGCAATTTAATCAAGCAATTTAATCAAGCAATTTAATCAAGCAAT 2468

QY 2336 -----ATTACATTAGACACATAGCCACCAATATGACCAATTTCGACTTCCGA 2386
 DB 2469 CAATGCTTATTATCTTAGCAACACCGGTAACAAAGGAGCTGACGCTTCCGA 2528
 QY 2387 CAATTCACTGCAACGCTGATATGCAAACTTGAACGTAATGCAATTTACGATTC 2446
 DB 2529 CAACGCTAGGCAAAAGCTAAGCCATTCGCACTCAACGCAATGCTCCCTAGCCATA 2588
 QY 2447 AGCTCAATTTCTTAAAAACAGCCATTTTCCACCAATTCAGGAGCAAAAGCCAC 2506
 DB 2589 GGCAGATTTCAATTTGAAAAACGCGCTTTACCGAAAAATAGCGGAGCAAGATAC 2648
 QY 2507 AACAGTAGCTGAAAAATGCACTTGAACATGCTAGGCAATCTACATTTGACAAATT 2566
 DB 2649 GGCATTTACCTTAAAAACAGCAATGACGCTGCTCCGCGGACGGAATTTAGCAATTT 2708
 QY 2567 AAGCTAAATACAGTACATCACTGTAATTCAGCTTAT----- 2606
 DB 2709 AATCTTGACACGCAACATTAACCTCACTCAATTCGCTATGACACAGATGCGGACGCG 2768
 QY 2607 -----TAGCTAGCTCAAAACATAGCCACG-----TCGCGCTTATAGAGACGGA 2653
 DB 2769 GCAACCGGACGATGCGGAAATGCGCGCGCGCGCTTCGCGCGCTTATTCGT 2828
 QY 2654 AACAGCGCAACATGCGGACACATGCTTTCAACATGACATGACATTAATGTAATGAG 2713
 DB 2829 TAGCGCGCAACTTGGAGAGATCCGCTTTGAACGCTGACGTAACGCAATTTGA 2888
 QY 2714 TGGGCAAGCACATTTCAATTTACTTCAATTTTATTTGGCTATTAAGCAATTAATA 2773
 DB 2889 CGGTACGAGCAATTCGCTTATATGCGMACTTGGGCTCCGCGGCAATTTGA 2948
 QY 2774 ATTATCCATGACGCTGAGCGGCAATTAATTAATTTCTTGGACACAGCGCAAGAC 2833
 DB 2949 GCTGGCGAAAGTTCCGAGGCACTTACCTTGGCTGTCAAAATACGCGCAAGAC 3008
 QY 2834 CGAAACCTTGAGCAATTAATTTGTTGTAAGCAAAATATCAACGCTTATCAATTA 2893
 DB 3009 CGTAAGCTCGAGCAATTTGACGCTTGAAGAAAGAAACACACGCGCTTCCGAAA 3068
 QY 2894 GCTCAATTTACTTTGAAATATGACAGCTTATGAGGAGCTTACCTTATTAATTAAT 2953
 DB 3069 TCTTAATTCACCTCGCAAAACGACAGCTGATCCGCGGCAATGCTTATGCTTAA 3128
 QY 2954 GAAGATGATGCGGAATTCGCTGATTAACCAATTAAGACAGCAAGTTGCAATGA 3013
 DB 3129 CCGAAAGAGCGGAGTTCCGCTGATTAATCCGCTTCAAGAAACAGAGCTTTCCGACA 3188
 QY 3014 TTTAGTAGAGGACGACGACGAAACGAACTTAGAGCCAAACAAATTTGAAACGAGTGC 3073
 DB 3189 ACTGCGAAGGCGGAGAAACAGAGGCGGCTTGAACGCAAAACAGGCAAACTTGGCGC 3248
 QY 3074 TAAAAACAAACAGGCTGAGCAAAATGCGGCTCAAGAAAGAGCGAGGAGCGCTTCC 3133
 DB 3249 CAAACAAACGCGGAAAAAGCAACGCGGAAACCTTGAACGCTTATTCGCGCGGCGC 3308
 QY 3134 TGAATACCTGCTGATCAAGCTGTTAAACGATTAAGACCAACAAAGCTGAATGAC 3193
 DB 3309 CAATGCCACGAAAGGCAAGAGCTTCCGCAACGCGCGCGGACGAGCGGCGGAAAA 3368
 QY 3194 TGCTGAACACAAAAAGTAAGGCAAAAAACAAAAATGCGCTCAAAAAAGACGTTT 3253
 DB 3369 TGCGGCAATTATGACGCGGAGGAAAGAAAAAGGCTGACGGAATTAAGACACCGC 3428
 QY 3254 TTTGATCCCTGCTTATCAAAAGCTGTTCCGATTAAGACGCACTTGAATTTCTGA 3313
 DB 3429 CTTGGCAAAACAGCGGAAAGGAAACCGCGCGGCTTACACGCGCTTCCCGCGCGC 3488
 QY 3314 TGCCCCACAGCAATGGAAGAAAGATGCTTAGCTCAAGAAAGAGCGGAAAAACACGCA 3373
 DB 3489 CCGGCGCGCGGAGATTTGGCGCAACGCGGCGCAACCGCAACCGCAACCGCGCGG-- 3546
 QY 3374 AAAAAAGACTTGATCAGCGCTTATTCAAAATGATGCTTATGAAATATCTGCAACAGT 3433

DB 3547 -----GACCTATCAGCGCTTATGCAATAGCGCTTGAAGTAATTTCCGCCAGCT 3589
 QY 3434 AATATGATGCTTCTGTTCAAGATGAATTAATGATGCTTTTGTATGATCAACCAATTC 3493
 DB 3600 CAACAGGCTTTTCCGCTTACAGACGAAATTTGACCGGCTGTTTCCGAAACCGCGCA 3659
 QY 3494 TCCGCTGAGCAAAATATGCAAGATTAAGCAAGCTTAATTTGATGCTTCCGCTG 3553
 DB 3660 GCGCTTTGACAGAGCGGATCCGGGACACAAACATACCGTTGCAAGATTTCCGCGC 3719
 QY 3554 TTTATGACGCAAAAGCACTTACTTAAATTTGGGGTGAAGAAACCTTGTAAATG 3613
 DB 3720 CTACCGCCAAC---AAACGACCTGCGCCAAATGCTATGCAAGAAACCTCGGACGCG 3776
 QY 3614 AGCATTTGGGCGAGTTTCTGCAATAGCGCTTCAATTAATCTTATGATGAACGTTAA 3673
 DB 3777 GCGGCTGGGATCTCTTTTTCGACAAACCGGAAACGCTTTCGACAGCGCATTCG 3826
 QY 3674 AATCAACGCACTTAACGATGATGTGGGTTTGCCCAATTAATGAGGCGATTTACA 3733
 DB 3837 CAATGCGCAACGCTTGCACAGGTCGTTTTCGGGCAATACGCAATCGGCAAGTTGA 3896
 QY 3734 ATTGCTGTAAAGCTGGGAAAGGGAATCACTGCGAGTAATGCTGAAGAACAAAGCG 3793
 DB 3897 CATCGGCAATGAGCGGCGCGGCTTTATGACGCGACGCTTCAACGCGCATKAGG 3956
 QY 3794 AAAAAATCAAGCAAAAGCATTAATTAATGCGCTGAATGCAAGTATCACTTCGTTAG 3853
 DB 3957 CAATTCGCGCGCGCGCTGCTGATTAAGGCAATTCAGGAAATACGCGCAAGTTCCG 4016
 QY 3854 GCAATTTGGGCAATTAACCTTATTTGAAGTAATGCTATTTATTAAGCTGAATTA 3913
 DB 4017 CGGATTTGGCAATGCAACCGCAATGCGGCAACGCTTATTTGTCGCAAAACGCGATTA 4076
 QY 3914 TCAATCTAGAGATGAGATGAAGTAAAGCGCTTACCTGATTAATGCTATTAATGCTG 3973
 DB 4077 CGGATGCAAAAGCTCAATGCAATGCGCAACCGGCGCTTGAATCAACGCTACCGCGG 4136
 QY 3974 CATGAGTTGATTAATTAATTTACTCCGACAGTAATTAATCAACGCTTATTTCTT 4033
 DB 4137 CATTAAGCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4196
 QY 4034 CGTCAATTAATGATGTTTCAACGCTTAACGTAACGTAACGTAACGTAACGTAACG 4093
 DB 4197 CCGTCTTATCGATGCGCTTCCGCAAGCTCCGCAAGCGCGCTATACCGCGTAT 4256
 QY 4094 GCAACCACTTTGACGCTTATTTGCAAAAGAGGAGGATTAAGGACAGAAATTTACA 4153
 DB 4257 GGGGAGGATTTGGCAAAACCGGCAAGTGCGGATGAGGAGGCTTAACGCAATTTCAA 4316
 QY 4154 TTTCCAAATTTCCGCTTATTTATCTCAAAATCTCAAGTTTCAACCTGCAACAGCAAA 4213
 DB 4317 TTTCAAGCTTCCCTCAGCGCTGCGCGCGGCAAGGCGCAATTTGAAGAGGACAG 4376
 QY 4214 TGTGGGCTGAATTTGGGCTATGTTGG 4241
 DB 4377 CCGGCGCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4404

RESULT 13

ACN43973

ID ACN40973 standard; DNA: 4404 BP.

XX

AC ACN40973:

XX

XX 27-0CT-2003 (revised)

DT 19-0UN-2003 (first entry)

XX

DE Prokaryotic essential gene #22610.

XX

KM Antisense; ds; prokaryotic essential gene; cell proliferation; drug design; gene.

XX
OS Neisseria gonorrhoeae.
XX
PN W020027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX P-PSDB; AB037103.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 28843; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation and the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway; (8)
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC ftp.wipo.int/pub/published_sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 4404 BP; 1271 A; 1170 C; 1094 G; 869 T; 0 U; 0 Other;
Query Match 30.0%; Score 1294.8; DB 7; Length 4404;
Best Local Similarity 59.1%; Pred. No. 3.4e-302;
Matches 2568; Conservative 0; Mismatches 1567; Indels 213; Gaps 12;
QY 86 TTTTAAACCGCTGATTCATTTAGGATAGTATTCGACAGGCGTGGGCGATCAACATTA 145
DB 78 TTACTTAGCCATATCCCTGCTGTGGCAATTCGCCCAAGCCCGGAGGAGACACATTA 137
QY 146 TTTTGGATTGATTACCAATATTATGATTTTCCGAGATTAAGGAGATTCACAGT 205

DB 138 TTTCCGACATCACTACCAATATCTATGCGACTTTGGCCGAAATTAAGGCGATTTGCACT 197
QY 206 TGGGGCTCAAAATATATAGGTTTATTAACAAACAGGCGATTTAGTGGCAATCAATGAC 265
DB 198 CGGGCGAAAGATATTAGGTTTACACAAAMAGGGAATTTGTTGGCAATGATGAC 257
QY 266 AAAAGCCCGATGATTTATTTCTGACTGTCAAGTAACGGGCGTGGACCTTGGTTGA 325
DB 258 GAAAGCCCGATGATTTATTTCTGCTGATCGGTGACGGGCGTGGCGCGG 317
QY 336 AAATCAATATATTGAGCGTGGCAATTAACGTAGATATACAGATTTGATTTGGTGC 385
DB 318 CGATCAATATATTGAGCGTGGCAATTAACGGGCGTGGCAATTAAGTTGATTTGGTGC 377
QY 336 AGAGGAAACACCCCGATCAACCGCTTTCTATTAAGATTTGAAACGAATTAACTA 445
DB 378 GAGGGAGAGCATCCCGATCGACCGCTTTCTTACCAATTTGTAAAGAAATATTA 437
QY 446 CAAAAAG--ATAATTTACATCTTATAGAGCAGATTACCATATTCACATTAATTA 502
DB 438 TAAAGCAGGAGCTAAGCGCATCTTATGCGCGGCAATTATGCGCGCTTGGACAA 497
QY 503 ATTGCTTACGAGCGCTCCATTTGATATGATTCGAAATTAATGAGCAGTACTATTC 562
DB 498 ATTGTCACGATGACGAACCTGTGATGATGACGTTATATGATGATGATGATGATG 557
QY 563 AGATGAAACAAATATTCAGAACGCTGTGATGCTGTGACGCGCATTTTGGCGAA 622
DB 558 TGATTTAATTAATATCTCTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 617
QY 623 TGATCAAGCAAAAGC-----GACCAAGTTCGGTGCATTAATTA 664
DB 618 TGATGAGAGCAACCAATACCGCGAAAGTTATATGCAAGGCAATTTCTTG 677
QY 665 TCTGACAGCTGCAATTAACCAATCAAGCTGACAGTATGATTTGATTTGG 724
DB 678 GCTGCTGTGGGATACCTTTGCAAAATGATCAGTGTGTCAGTCACTTGA 737
QY 725 AGGGGATTTGTAAGGGGAGAAATGAGTCAATTCAGTTCAGGCTCAAGGGGGA 784
DB 738 TAGGAAAAATTAATCATGACCCATATGCTTTTTCACAGAGGCTCATTTGGGA 797
QY 785 CAGTGTTCCTGATGATTTATTTATGATGCTGAAAAACAAATGTTAATATGAGAT 844
DB 798 CAGTGTTCCTGATGATTTATTTATGATGCTGAAAAACAAATGTTAATATGAGAT 857
QY 845 ATTACGAAAGGACCTTTTGAAGGCAAAATGCTTCAATTTGCTGCAATTC 904
DB 858 ATTGCAACAGGCAACCCCTATATGAGAAACCAATGCTTCAAGTATGTAAGA 917
QY 905 TATTT---TGATGAAATTTTGAAGATTTTACATCACTTACACCGAGCTGG 961
DB 918 TTGCTTCAATATGAAATTTTCTGAGATACCATTCAGATTTCAAGAACATCA 977
QY 962 TATGAGTGTACATATGTAAGTAATGATGCTGAGGCTCAT-----ACT 1013
DB 978 AATGGAATATCTTTTAAACCAATATATGCGGAGAAATATGATCCCAACA 1037
QY 1014 CAGAAATCAGATTAACATCAAGAAATTAATTAAG-----TTACCAATATGAG 1063
DB 1038 TAAACATATCTTACCTTATGATTAACACAGAACCGTTCAATTTATTTTC 1097
QY 1064 TTTACCTTGAAGAGATTAAGTTCAATATCTTATATGACGAGACCTTAATTTA 1123
DB 1098 TTTATTCAGACGACAGAGAACCTTTTATCTATGTA---GATGGGTCACAGTTA 1154
QY 1124 TTCTCAGCTTAAATATGAGAAAGCTATATTTATGATCAAAAACAGAGATCAT 1183
DB 1155 TCGAACCAAGTGAATTAATGAGAAATATTTCTTATTTGCAAGGAAAGTGAAT 1214
QY 1184 AATCTTGCATCTGACATTAACAAAGGGGCGGTGCTTTTGTGAGGATATTTTC 1243

Db 1215 GATACCTACCAACATCAACCAAGCGCGGCGGTTGATTTGAGGGTAATTAC 1274
Qy 1244 AGATATCCAAATTCMAACAACTTGGCAAGAGCTGGCATACATGTAAGAAAATG 1303
Db 1275 GGTCTGGCTAAACCAACGAAACGGCGAGCGGGGTTTCATACATGATGAGCG 1334
Qy 1304 CACCGTACTTGGAAAGTAATGCGTGGAAACATGATCGACTTTCMAATTTGTAAG 1363
Db 1335 TACCCTACTTGGAAAGTAATGCGCTGGCAACACCGCTTCCTCCMAATTCGCAAG 1394
Qy 1364 AACATTCAGCTTCAAGCCAAAGGGAAATTAAGTTTCATGATGAGCGATGTA 1423
Db 1395 CAGCGTCTGGTTCAAGCAAAAGGGAAACCAAGGCTCGGTGAGGGGCGAGCGTAA 1454
Qy 1424 AGTCAATTTGGAGCAGCGCAACATCAAGCGCAAAACAGCTTTAGTAAATGG 1483
Db 1455 AGTCATCTTGAATCAGCGCGACATCAAGCGCAAAACAGCTTTAGTAAATGG 1514
Qy 1484 CTGGTTAGCGCAGAGGACCTGTCATTAACATTAACATTAACATTAACATTA 1543
Db 1515 CTGGTACGCGAGGGAGAGCGTGAACATGATGCGATATACATTAACATTAAC 1574
Qy 1544 ATTATATTCGCTTCCTGGTGGTGGCTTGAATCAAGCGCATTCATTAACCTTAA 1603
Db 1575 ACTCATTCGCTTCCTGGCGGAGCTTGGATTTGAAGCGCATTCCTTCCTTCCA 1634
Qy 1604 AGCTATCCAAATTCAGCAGAGGGCGCATTAATTTGAACCAATTAACATTAAC 1663
Db 1635 CCGCATTCAAATTCAGATGAAAGGGCGATGATGTCAACCAACCAAGCAAGATC 1694
Qy 1664 TAAATCATTAATTCAGCAAGCAAGCATTTCTTCACTTAATGAATTAATTA 1723
Db 1695 CACGTTACATTAACAGCAATTA-----GATATTCATCAACCGGCAATTAAC 1748
Qy 1724 ACTGATTAACAGAAAGTAATTCCTTCAACAGCTTGGTGGCAACAGATTAATTA 1783
Db 1749 CTGGATTCAGAAAGTAATTCCTTCAACAGCTTGGTGGCAAGATTAATTA 1808
Qy 1784 ACATATTCGCGATTAACCTTATTAATTAACCAACCAAGATTCATTAATTA 1843
Db 1809 AACCAAGCGCGCTCATCTGAATTCACACCGAAGACAGATTCATTAATTA 1868
Qy 1844 TTCAGTGTGACAAATTTAAAGGCGATTTACCAACCAAGATTAATTTTAC 1903
Db 1869 TTCGCGCGAAACAAATTTAAAGCGCATTAACCAACCAAGATTTTAC 1928
Qy 1904 CGGTAGACGACACCGCAAGCTTCAATTAATTAATTAACGTTTGTCAAGATGAAG 1963
Db 1929 CGGACAGCGACACCGCAAGCTTCAATTAATTAATTAAGCGGATGAAG 1988
Qy 1964 TATACCAAGCGGAATTTGTGTGGATTCAGATTTGATCAACCTTAATTAAGTGA 2023
Db 1989 TATCCCAAGAGGAATTCGTGTGGCAACCAATTTGATCAACCTTAATTAAGTGA 2048
Qy 2024 AAACCTTCAAATTTAAAGCGGAAGTGGCTGCTTCGCAATTTTCTCAATTAAGG 2083
Db 2049 AAACCTTCATTAACAGGCGGAGCAAGGCTGCTTCGCAATTTTCTCAATTAAGG 2108
Qy 2084 AAATTCAGACGTCAGCAATTAATTCGAATTCGAATTTGTGTTCGAATTAAG 2143
Db 2109 CGATTCGATTTAAGCAATTCAGCCCAAGCTTTGCTGTCACCGATCAAGCA 2168
Qy 2144 TACATTTGACCGCTTCAATTCAGATTCAGATTTAAGCTTTCGAATTAAGTGA 2203
Db 2169 CACATCTGTACAGTTGCGAGCTGAGCGGTGCAAGTTGTACGAAACCAATTA 2228
Qy 2204 CGATCAAAAGTATTAATTCATTAACCAAAACCAATTCATTAATTAATTAAC 2263
Db 2229 CGAGATTAAGTATTCCTTCAATTCAGCAAGCAAGATTCAGCAATTCAGCTTCA 2288
Qy 2264 TGATATTCAGCGCGATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 2323
Db 2289 CGATTCAGCTTAAATTCAGAGATTCAGCAACCTCAACCGGCAATTCATTAAGTGA 2348
Qy 2324 AAATCAGCA----- 2335
Db 2349 CGGACACCGCATTAACGTTTACGCGCAACCGCAACCAAGCAACCTTACGCTGT 2408
Qy 2336 ----- 2335
Db 2409 GGGCAATTCGCAACATTTAATCAAGCATTAACGCAATTAACGCAACATTCGCTTCA 2468
Qy 2336 -----ATTATCATTAAGCAATTCGCAACCAATTAAGCAATTTGATTCGA 2386
Db 2469 CAATCTTCATTAATTAAGCAACAGCGCTTAACCAACGCGCTTGAAGCTTTTCA 2528
Qy 2387 CAATTCATTCAGCGCTGATTAAGCAATTCGAACGCTTATGCAATTAAGCAATTC 2446
Db 2529 CAACCTTAAGCAACGTAAGCAATTCGCACTCAAGCGCAATTCCTTACGATTA 2588
Qy 2447 AGCTCAATTTCTTTAAACAGCAATTTTTCGCAACCAATTCAGGACCAAGGCA 2506
Db 2589 GCGATTCATTCATTTGAAACAGCGCTTACCGGAAATCAAGCGCGCAAGGATAC 2648
Qy 2507 AACATGACGTTGGAATTCGCACTTGGACATGCGCTTACCAATTCATTCAGCAAT 2566
Db 2649 GCGATTCATTAATTAAGCAAGCAATTCGCTGCGCTGCGCAAGATTAAGCAAT 2708
Qy 2567 AACGTAATTAACGATACATTCATTAATTCAGCTTA----- 2606
Db 2709 AACCTTCAACAGCGCAATTCATTAATTCAGCTTAATTCAGCTTAATTCAG 2768
Qy 2607 -----TCAGTACGCAACATTAAGCGCA-----TCGCTTCATTAAGCGCA 2653
Db 2769 GCAACCGCGAGTGCAGATTCGCGCGCGCGCGCTTGCAGCGCTTCATTAATTC 2825
Qy 2654 AACAGCGCAATTCGCGCAACATTCGTTCAACATTAACAGTAATTCGTAATTCAG 2713
Db 2829 TACCGCGCAATTCGCGCAATTCGCTTTCAGCGCTTAACAGCTTAAGCGCAATTC 2888
Qy 2714 TGGCAAGCGCAATTCATTAATTCATTCATTTTGGCTTAACCAAGCAATTAATTA 2773
Db 2889 CGGTCAAGCAATTCGCTTATTCGCAATTCGCTTACCGCGCAATTCGCAATTC 2948
Qy 2774 ATTATCATACCGTGAAGCGATTAATTCATTCATTCGCTTTCGCAACAGCAATTC 2833
Db 2849 GCTGCGCAAGGTTCCAGAGCACTTACCTTGGCTTCAACATTAACGCAACCAAC 3008
Qy 2834 CGAAACCTTCAGCAATTCGTTGTTGAAGCAAGATTAATTCAGCTTATTCAGATTA 2893
Db 3009 CGTATCTTCAGCAATTCAGCTTATTCAGCTTATTCAGCTTATTCAGCTTATTC 3068
Qy 2894 GCTCAATTTTACTTGAATTAACAGCTTATTCAGCTTATTCAGCTTATTCAGCTT 2953
Db 3069 TCTTATTTTCACTTCGCAACCAAGCAATTCAGCTTATTCAGCTTATTCAGCTT 3128
Qy 2954 GAGGATTCAGCAATTCGCTTTCATTAACCAATTAAGCAAGATTCAGCAATTA 3013
Db 3129 CCGCAAGAGCGCAATTCGCTTTCATTAACCAAGCTTTCAGCAATTCAGCTTTC 3188
Qy 3014 TTTATTAAGAGCAAGCAAGCAAGCAATTCAGCAATTCAGCAATTCAGCAATTC 3073
Db 3189 ACTCGCAAGCGGAGCAAGCAAGCGCTTTCAGCAATTCAGCAATTCAGCAATTC 3248
Qy 3074 TAAACCAACAGGTAGCGCAAGCTTTCAGCAATTCAGCAATTCAGCAATTCAG 3133
Db 3249 CAAGCAAGCGGAGCAAGCAAGCGCTTTCAGCAATTCAGCAATTCAGCAATTC 3308
Qy 3134 TGTATTCCTGCTATCAAGCTTTCAGCAATTCAGCAATTCAGCAATTCAGCAATTC 3193
Db 3309 CAATTCAGCAAGCGGAGCAAGCGCTTTCAGCAATTCAGCAATTCAGCAATTC 3368
Qy 3194 TGTGAAACCAAGGTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3253
Db 3369 TGGCGCATTAAGCGCGAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3428

QY 3254 TTCTATCCCTTCTGATCAAGCTGCTGCTATGAAAGCCGCACTTGAGTTATGA 3313
 Db 3429 CTTCGGGAAAGAGCGGCAAGCGAAACCGCGCGCTTACACCGCTTCCCGCGCGG 3488
 QY 3314 TGCCCCAGCAATCGGAAAAAGATGCTGACTCAAGAAAGCGGAAAAACAAGCA 3373
 Db 3489 CCGCGCGCGCGGATTTGCGGCAACCGGACCCCAACCGCAACCCCAACCGGCGC-- 3546
 QY 3374 ACAAAGCTTGAATGACCGCTTATTCATTAAGGCTTATCAAGATTATCGAACGT 3433
 Db 3547 -----GACCTGATCAGCGGTATTCGCAATGCGGTTTAAAGTAAATTTCCGCGCT 3599
 QY 3434 AAATAGTATGCTTCTGCTGATCAAGTATGATGCTGCTTCTTGTGATGATCAAGCAATC 3493
 Db 3600 CAACAGCGTTTTCGCTGATCAAGCAATGACCGGCTGTTGCGGAAAGCGCGCA 3659
 QY 3494 TGCCGCTGGAATATGCGACAGATTAAGACCTGATGATTCGATGCTTCCGTCG 3553
 Db 3660 CGCGGTTTGAACAAGCGGCAATCGGGAACCAACACCTGCGCAAGATTTCCGCGC 3719
 QY 3554 TTATCGAGCAAGAAAGCACTTACCTAAATTTGGGTCGCAAAAAAGCTTACGTAATG 3613
 Db 3720 CTACCGCCAAAC---AAACCGACTGCGCAATGATGACGAAACCTGCGACGCG 3776
 QY 3614 ACGAATTTGGGCGAGTTTCTGCAATGACCGTTCAGATTAATACCTTTGATGAACAGGTTAA 3673
 Db 3777 GCGGCTGCGCATCTGTTTTCGCAACAACCGGCAACCTTTCAGACAGCGCATCG 3836
 QY 3674 AAATCAGCGACATTAACGATGATGCGGCTTTTGCCCAATATCAATGCGGCGATTTACA 3733
 Db 3837 CAACCTGCGACGCTTGGCCACGCTCCCTTTTGGGCAATACGCGCATTCGCGAGTTCGA 3896
 QY 3734 ATTGCTGTAACGCTGGGAAAGGCAATGAGTGGCAATTAAGCTGAAAGCAACCG 3793
 Db 3897 CATCGCATCAGCGCGCGCGGCTTTTAAAGCGGCAAGCTTTCAGCGCATCAGAG 3956
 QY 3794 AAAAATTCATCGAAGCAATTAATTAATGAGTGAATGATTAATGATTCGTTTAAAG 3853
 Db 3957 CAATATCCCGCGCGCGCTGCTGATTAAGGCAATTAAGGCAATTAAGGCAATTAAG 4016
 QY 3854 GCAATTTGGGCACTGAGCTTATTTTGAATGATGCTTATTTATGAAAGTGAATTA 3913
 Db 4017 CGATTCGCGCATCGAAGCGCAATCGCGCAACGCGCTATTTGTCGCAAAAGCGATTA 4076
 QY 3914 TCATCTGAGGAGGAGGAGGAGGAAAGCGCTTACCTGCAATTAATGCTTAATGCTG 3973
 Db 4077 CCGATGAGAAAGCGCAATATCGCAACCGCGGCTTTCATTAACCGCTTATTTGAG 4136
 QY 3974 CATTCGATGATTAATTAATTAATTCGCAAGATTAATGAGGCTTATTTTCTT 4033
 Db 4137 CATTAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4196
 QY 4034 CGTCAATTTATGTTGATTTTCAACGCTTACGTAAGCAACGCTTAAATCTCAAGGTT 4093
 Db 4197 CCGTCTCTATACCGATGCGCTTCCGCAAAAGTTCGAAACGCGGCTTAATTCGCGATTT 4256
 QY 4094 GCAACCAATTTGAGCTTATTTGAGGCAAAAGAGTGGGCTTAAGGCAAGATTTTACA 4153
 Db 4257 GCGCGAGATTTTGGCAAAACCGCGCATGCGGAGATGCGGCTTAAGGCTTAAGG 4316
 QY 4154 TTTCCAATTTCCGCTTTTATCTCAAAATCTCAAGTTCAACTCGGCAACAGCAAAA 4213
 Db 4317 TTTCAGCGCTGCTCCACGCTGCGCGCGCAAGGGGCGCAATGGAAGCGACAG 4376
 QY 4214 TGTGGCGGTGAATTTGGGCTTGTGTTGG 4241
 Db 4377 CGCGGCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4404

RESULT 14
 AAZ38918
 ID AAZ38918 standard; DNA; 4374 BP.

AC AAZ38918;
 XX 21-FEB-2000 (first entry)
 DT
 XX
 DE Neisseria meningitidis strain H44/76 BASB006 nucleotide sequence.
 XX
 XX Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
 KM vaccine; antibiotic; upper respiratory tract infection; meningitis;
 KM invasive bacterial disease; bacteraemia; screening; antibacterial; ss.
 XX
 OS Neisseria meningitidis.
 XX
 XX
 FH Key Location/Qualifiers
 FT 1..4374
 FT CDS /*tag= a
 FT /product= "BASB006"
 FN W0955873.f2.
 PD 04-NOV-1999.
 PF 20-APR-1999; 99WO-EP002766.
 PR 24-APR-1998; 98GB-00008866.
 PA (SMIX) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 P1 Thomard J;
 P2 WPI; 2000-052810/04.
 DR P-PSDB; AAT56622.
 XX
 PT Novel polynucleotides and polypeptides from Neisseria meningitidis used to
 PT prepare vaccines against bacterial infections.
 PS Claim 11, Page 91-93; 103pp; English.
 XX
 XX The present sequence encodes BASB006 isolated from Neisseria meningitidis
 CC strain H44/76. BASB006 polynucleotides and polypeptides may be employed
 CC as research reagents and material for the discovery of treatments and
 CC diagnostics for diseases, particularly human diseases. They can be used
 CC for diagnosis of disease, staging of disease, or determining response of
 CC an infectious organism to drugs. The polynucleotides may be used as a
 CC source for hybridisation probes, and for screening of genetic mutations,
 CC serotype, organism or strain identification, identification of mutation
 CC in BASB003 sequences, and as components of arrays which are useful for
 CC diagnostic and prognostic purposes. The polypeptides can be used to
 CC produce antibodies. The polypeptides can also be used in vaccine
 CC formulations, and to identify agonists and antagonists. The polypeptides,
 CC antibodies, agonists and antagonists (which are bacteriostatic) are used
 CC for the treatment and prevention of diseases such as upper respiratory
 CC tract infection, invasive bacterial diseases such as bacteraemia and
 CC meningitis, and for the development and screening of antibacterial drugs.
 CC They are also used in the prevention of adhesion of bacteria to
 CC eukaryotic matrix proteins on in-dwelling devices, or to extracellular
 CC proteins on wounds, and to prevent tissue damage and/or block the normal
 CC progression of pathogenesis in infections initiated other than by the
 CC implantation of in-dwelling devices or by other surgical techniques
 XX
 SQ Sequence 4374 BP; 1269 A; 1158 C; 1066 G; 881 T; 0 U; 0 Other;
 Query Match 29.0%; Score 1252.6; DB 3; Length 4374;
 Best Local Similarity 58.8%; Pred. No. 5.3e-282;
 Matches 2567; Conservative 0; Mismatches 1534; Indels 264; Gaps 13;
 QY 81 CTTAATTTTAAACCGCTTCAATTCATTAGGATAGATCGAAGCGTGGCTGCTAC 140
 Db 73 CCGCTTACTTACGCAATATGCTGCTGCTGCTGCTTCCCAAGCTGCGGCGGACAC 132
 QY 141 ACTAATTTGGATGATTAACCAATATTAATGCTGATTTTGGCGAATTAAGGCAAGTTC 200
 Db 133 ACTAATTTGGATGATTAACCAATATTAATGCTGATTTTGGCGAATTAAGGCAAGTTC 192

QY 201 ACAGTGGGCTCAAAATATTAGTTATTAACAACAGGCGAATTAGTTGGCAGATCA 260
DB 193 GCAGTGGGGCGAAGGATTTAGTTACACAAAAAGGGAAGTTGGTGGGAATCA 252
QY 261 ATGACAAAAGCCCGATGATTTTCTGTAGTGCATGAAGCGGTGGACCTTG 320
DB 253 ATGACAAAAGCCCGATGATTTTCTGTAGTGCATGAAGCGGTGGACCTTG 312
QY 321 GTTGAATGCAATATTTGAGCGGCACTAAGTATGATGATGATTT 380
DB 313 GTGGGCGATCAATATTTGAGCGGCACTAAGTATGATGATGATTT 372
QY 381 GGTGACAGGAGAAACACCCGATCAACATCTTTTACTTATAGTTGTAACGAAAT 440
DB 373 GTGGCGGAGAGAAATCCCGATCAACATCTTTTACTTATAGTTGTAACGAAAT 432
QY 441 AACTCAAAAAAG--ATATTTACATCTTATGAGAGCATTAACATATCAAGATTA 497
DB 433 AATTATTAAGGAGGAGCACTAAGGCGATCTTATGTTGGCGATATCAATGCGCGTTTA 492
QY 498 CATTAATGCTTACAGAGGCGCTCAATGATTAAGTCTTGAAATGATGAGAGTACT 557
DB 493 CATTAATTTGTACAGATGACAGAACTGTTGAATGACCATTTATGATGAGCGGAGAA 552
QY 558 TATTCAGATGAGACAAATATCCAGACGTGTGCTATCGGCTTGGACGCGAGTTTGG 617
DB 553 TATTCAGATGAGAAATTAATCCGATGACGCTGTGCTATGAGGACAGGAGAAATTTGG 612
QY 618 CGAAATGATCAAGCAAAAGC-----GACCAAGTTGGCGGTGAT 659
DB 613 CGATCTGATGAGATGAGCCCAATTAACCGGAAAGTTCAATATTTGCAAGGCGCTAT 672
QY 660 CATTAATGAGCGGAGATGACAGCAATCAGGTGGAGAGATGATGATTTGCTAT 719
DB 673 TCTTGCTGTGTGGTGGCAATCTTTGCAAAATGATGAGTGGTGGCAAGTAC 732
QY 720 TTGGGAGCGAGATCTGTAAGCGGAGAAATGATGATGATGATGATGATGATGATGAT 779
DB 733 TTGAGTGTGAGAAATTTAAACATGACCATTTGTTTAAACAGAGGCGTCAAT 792
QY 780 GGGGACAGTGTCTTCGATGTTTATTAATGATGCTGAAACAAATATGTTATTTAT 839
DB 793 GGGCAGAGTGTCTCAACATGTTATCTATGATGCTGAAACAAATATGTTATTTAT 852
QY 840 GGGGATTTAGGAGAGCAACCTTTGAGGAGAAATGATGATGATGATGATGATGATGAT 899
DB 853 GGGGATTTAGGAGAGCAACCTTTGAGGAGAAATGATGATGATGATGATGATGATGAT 912
QY 900 AATCTTATTT--TGATGAATTTTGAAGAGATTTAATCATCATCTTTACACCGA 956
DB 913 AATGATTTGTTCTATGATGAAATCTTTGCTGAGATACCATTTCAATTTCTACAGAAC 972
QY 957 GCTGATATGAGTGTACATTTAGTGAATGATATGCTGAGGCGTCTATACAG 1016
DB 973 CGTCAAAATGAGAAATCTTTTACAGAGATTAATATGAGCAAGAGAAATCAATGCC 1032
QY 1017 AATATCAGGA-----ATACATCAGAAATTAATATGCTTATGACCAAT 1058
DB 1033 AATATATACCAATTTCTGCTCAATATTAATTAATTAATTAATTAATTTATTTAT 1092
QY 1059 ATGAGTTTACTTTGAAAGAGAGATTAAGTTCAATATCTGATATGACGAGCTTAT 1118
DB 1093 GTTTCTTTATCCGAGAGAGAGAACTGTTTATCATGTGCA---GGTGGTGTCAAC 1149
QY 1119 ATTTATTTCTGACGTTTAAACATGAGAGAAACGCTATATTTTATGATGATTAACAGGA 1178
DB 1150 AGTATGACCCAGCTGATATATGAGAGAAATTTCTTATATGACAGAGAGAAAGGC 1209
QY 1179 TCATTATCTGCACTGACATTAACCAAGGAGGCGGTGCTTTATTTTGGAGGTAT 1238
DB 1210 GAATGATATCTTACAGAGACATCAATCAAGGTGCTGAGAGATTAATTTTCCAGAGAT 1269
QY 1239 TTACAGTATCTCCAAATTTCAACAACTTGGCAGAGAGCTGGCATACATTAATGATA 1298

DB 1270 TTACGCTCTGCTGAAATTAAGAAACGTTGCAAGGTGGCGCTTCAATTCAGTAA 1329
QY 1299 AATGACCCCTTCTCTGAAAGTAAATGGGCGGAGAAATGATGATGATGATGATGATGAT 1358
DB 1330 GACATGCTCTCTGAAAGTAAATGGGCGGAGAAACGCTGTCGAAATTCGAC 1389
QY 1359 AAGGACATTTGACGCTTCAAGCCAAAGGAGAAATTAAGGCTTCAAGCGTACGAT 1418
DB 1390 AAGGACGCTTCAAGCTTCAAGCCAAAGGAGAAACGAGCTGATCAAGCGTGGCAG 1449
QY 1419 GGTAAAGTATTTGAGACGACGAGGAGATCAAGGATCAAGGATCAAGGCTTTAGTAA 1478
DB 1450 GGTACAGTATTTGATCAGAGGAGAGATTAAGGAGAAACGAGCTTTAGTAA 1509
QY 1479 ATTTGCTGTTGAGGAGGAGGAGGAGCTTCAATTAAGATGATTAACATTTGATCC 1538
DB 1510 ATGCGCTTGGTGGAGGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGAT 1569
QY 1539 GATTAATTTATTTGCGCTTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1598
DB 1570 GACAAATCTATTTGCGCTTTGCGGCGAGCGTTGATTAACGAGGCAATGCGCTTGG 1629
QY 1599 TTAAACGATTCAAATATGAGAGGAGGAGATGATGATGATGATGATGATGATGATGAT 1658
DB 1630 TTCCACGCTTATCAAAATCCGATGAGAGGAGGAGATGATGATGATGATGATGATGAT 1689
QY 1659 GCGCTATGCTATCTTACTGGAGAGAAAGATGATGATGATGATGATGATGATGATGAT 1718
DB 1690 GAATCAACCTTATCTTACAGGATTAAGATTTG-----CTACAGCCGCAATAC 1743
QY 1719 AATTAATCTGATTAAGAAAGAAATTTGCTTACAGCGTTGTTGGAGAGATTA 1778
DB 1744 AATGATCTGATTAAGAAAGAAATTTGCTTACAGCGTTGTTGGAGAGATTA 1803
QY 1779 AATTAACATTTGAGGAGATTAACCTTATTTATTAACCAACAGAGATGATGATTTG 1838
DB 1804 ACCAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1863
QY 1839 CTACTTCAAGGTTGATCAAAATTTAAAGGCGATTAACCAAGGAGGAGGAGGAGGAGGAG 1898
DB 1864 CTGCTTTCCGCGGAGAAATTTAAAGGCGATTAACCAAGGAGGAGGAGGAGGAGGAGGAG 1923
QY 1899 TTACGAGTACAGGAGCAACCGGAGCGCTTCAATCAATTTAAATTAACGTTGGTACAG 1958
DB 1924 TTACGAGTACAGGAGCAACCGGAGCGCTTCAATCAATTTAAATTAACGTTGGTACAG 1983
QY 1959 GAGGATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2018
DB 1984 GAGGATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2043
QY 2019 GGTGAAATCTTCAATTTAAAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2078
DB 2044 GGGGAAATCTTCAATTTAAAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2103
QY 2079 GAGGAAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2138
DB 2104 AAGGAGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2163
QY 2139 GAAATATCATTTGACGCTTCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2198
DB 2164 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2223
QY 2199 TTACGAGATCAAAAGTATTAATTTCAACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2258
DB 2224 ATTTACGAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2283
QY 2259 TTACGAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2318
DB 2284 CTGCGAGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2343
QY 2319 TTAAACATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2378

Db 2344 GCAGATGGGATACAGCTTATACAGTCAAGCCACAAAGCCAAAGGCACTTAGC 2403
 QY 2379 ----- 2378
 Db 2404 CTGCTGGGCAATGCCCAGACATTATATACAGCCACATTAAAGGCAACATCGGCT 2463
 QY 2379 -----CTT 2381
 Db 2464 TCGGCAATGCTTCAATTATCTAAGGACCAAGCCGTACAAAGGCAAGTCTAGGCTT 2523
 QY 2382 TCCGACATTTCACTGCAAGCTGATATATGCAAACTTGAATATGTCATTTAAG 2441
 Db 2524 TCCGGGCAAGCTAAGGCAAGCTAAGCCATTCCGACTCAAGGATATGTCCTTAGCC 2583
 QY 2442 GATTGAGCTCAATTTCTTTAAAAAAGCCATTCTTGCACCAATTCAGGGAGACAA 2501
 Db 2584 GATAGGCAAGTATTCATTGTAAGACAGCCGCTTTACCCGGAATATAGGCGGGAAG 2643
 QY 2502 GGCACACAGTGAAGTGGAAATGGAAGCTTGGACATGCTAGGATATACATTGCA 2561
 Db 2644 GATAGGCAATTCACCTTAAAGACAGCCGATGAGCTGCGCTAGGCAAGCAATTAGG 2733
 QY 2562 AATTACGCTAATTAACAGTACATCACTTAATTCAGCTTAAT 2606
 Db 2704 AATTAACTTGAACAGCCACATTACATCTAATTCGCTATAGCCAGATAGGCA 2763
 QY 2607 -----TAGCTAGCTAAACATAGCCAG-----TCGCGCT 2639
 Db 2764 GGGGCGCAACCGGAGTGCAGACATGCGCGCGCGCGCTTTCGCGCGCTTCGCGCG 2823
 QY 2640 TCATTAGACAGGAAACAGCCACATCGAGAGACATGCTTAAACATTAGACAT 2699
 Db 2824 TCCCTATATTCGTTACCGGCACTGATGATGATCCGTTTCAACAGCTAGCGTA 2883
 QY 2700 AATGTAATATGATGAGGCAAGGACATTCATTAATTCATCTTATTTGGCTATTA 2759
 Db 2884 AAGCGCAATTTGAACGCTAGGGAAGATTCGCTTTATGTCGAACTTCGCGCTAC 2943
 QY 2760 AGCGATAATTAATTAATTCATGACGCTGAGGCGATTCATATTAATTCGTCGAC 2819
 Db 2944 AGCGCAAAATTGAACCTGCGGAAAGTTCGGAAGCACTTACCTTGGCTCAACAA 3003
 QY 2820 ACAGGCAAGAACCCGGAACCTTGAACATTAATTGTTGTTGAAGCAAAATATCA 2879
 Db 3004 ACCGCAAGCAAGCTGCAAGCTTGAACATTTGACGTAATGTAAGAAAGAAACAA 3063
 QY 2880 CCGTTATCAGATTAAGCTCAATTTACTTTAGAAATGAACAAGTTAGTCAAGTGCAT 2939
 Db 3064 CCGCTGTCGAAAACCTTAAATTTCACTTGAACAAACGTCATGCGCGCTG 3123
 QY 2940 CGTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2999
 Db 3124 CGTTACCACTATCCGCAAGAGCGGAGTTCCGCTGATATCCGCTCAAAACAA 3183
 QY 3000 GAATGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3059
 Db 3184 GAGCTTTCGCAAACTGCGCAAGC-----AAGACGAAA 3222
 QY 3060 GTTGAACGACTGCTTAAACAAACAGTGAAGCCAAAGTGGCTCAAGAAAGACG 3119
 Db 3222 CAGCGGAAAGAAACAAACGCGCAAGCTTGAACGCTGATGCGCGCGCGGCAATGCC 3282
 QY 3120 AGAGCAAGCTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCTTAAAGCCAA 3179
 Db 3283 GTCAAAAG---ACAAAGAGCTTGCAGAACCGGCTCGGCAAGCGGGAAGAAATGTC 3339
 QY 3180 CAAGTGAATGCTGCTGAACAAACAAAGAAAGTAAAGCAAAAGTGGCTCA 3239
 Db 3340 GGCATTATGAGCGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3399
 QY 3240 AAAAGCAAGCTGTTTCTGATCCCTGCTTGAATCAAGCTGTTGCTTGAATGAAGCC 3299
 Db 3400 GCGAAACAGCGGAGAGCGGAAACCGCGCGGCTACACCGCTTCCCGCGCGCGCGCG 3459

QY 3300 CTTAGGTTATGATGCCCAACAGCAATCGGAAAGATGCTTAGCTAAGAGAGAGCG 3359
 Db 3460 GCGCGCGGAGATTTGCGCGAAGTGAACCCAAACGCGAGCCGCAAC----- 3506
 QY 3360 GAAAAACAGCAACAAAAAGACTTATAGCGCTTATCAAAATAGTGGTTATCAAA 3419
 Db 3507 -----GCAAGCGCAAGCTGATCAAGCGTTATAGCAATACCGGTTAGATGA 3552
 QY 3420 TTATCTGCAAGTAATAGTATGCTTCTGTTCAAGATAGATAGATAGATAGATAG 3479
 Db 3553 TTTTCGCAAGCTCAACAGGCTTTGCGCTTAAAGAGAAATTAAGCCGATTTGCG 3612
 QY 3480 GATCAAGCAATCTCCGCTGACAAATATGCAACAGATTAAGAACGCTATGATTC 3539
 Db 3613 GAAAGAGCGGCAAGCCGCTTTGCAAGAGGCAATCGGCAACCAAAACATACGCTTG 3672
 QY 3540 GATGCTTCGCTCTTATAGCAAGCAAAAGCAATTAAGCTTAATTTGGGATGCAAAA 3599
 Db 3673 CAAGATTCGCGCCCTACCGCCAC---AAACGACCTTGGCAAAATGGATATGCAAAA 3729
 QY 3600 GCTTAGCTAAGAGCAATGAGGAGATTTCTGCAATAGCCGTCAGATTAATACCTT 3659
 Db 3730 AACTCGCAAGCGGCGGCTGCAATCTGTTTTCGCAACCGGAAACGCAATCTTC 3789
 QY 3660 GATGAACAGTTAAATATCAAGCAATTAAGCATGATGATGCTTGGCCATATCA 3719
 Db 3790 GAGCAAGGCAATCGGCAACTCGCAAGCTTGCACAGCGCGGCTTTCGCGCAATACG 3849
 QY 3720 TGGGAGCATTAATTAATTTGCTTAAAGTGAAGCGGCAATGATGATGATTAATAGCT 3779
 Db 3850 ATGCAAGATTTATCATCGCATCAAGCGGCGGCTTTTGAAGCGGCAACCTTTCA 3909
 QY 3780 GAAAGCAAAAGCGGAAATTCATGCAAAAGCGATTAATTAAGCGGATTAAT 3839
 Db 3910 GAGGCAATGAGGCAAAATCCGCGCGGCTGCTGATTAAGGATTCAGCAACATC 3969
 QY 3840 CAGTTCGTTAGGCAATTTGGGCAATTCAGCTTATTTGAAGTAAATGCAATTTAT 3899
 Db 3970 CCGCGGCTTTCGCGGATTCGCAATCGAACCGCAATCGGCGCAACCGCTATTCGTC 4029
 QY 3900 GAACTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3959
 Db 4030 CAAAAAGGAGATTCAGCTACGAAACGTCATATGCGCAACCCCGGCTTGCATTAAC 4089
 QY 3960 CGCTAATAGCTGCAATTCAGTATTAATTAATTAATTAATTAATTAATTAATTA 4019
 Db 4090 CGCTACCGCGGCAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTA 4149
 QY 4150 AGCCCTATTTGAGCCCTGCTTATACGATCGCGCTTGGGGAAGTCCGAACAGCGTC 4209
 Db 4080 AATTCAGGCTTGAACCAACCTTGAAGCTTATTTGCAAAAGAAAGTGGATTAAG 4139
 QY 4210 AATACCGCGTATGCTCAGATTTTCGCAAAACCGGCAATGCGGATGCGGATTAAC 4269
 Db 4140 GCGAATTTTCAATTTCCAAATTCGCTTTATTTCAAAATCAAGCTTCAACTC 4199
 QY 4270 GCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4329
 Db 4200 GCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4244
 QY 4330 GAGGCGCAACAGCGGCGCATCAATTAAGCTACGCTGTA 4374

RESULT 15

AAZ12251

ID AAZ12251 standard; DNA; 4374 BP.

XX AAZ12251;

AC 08-OCT-1999 (first entry)

XX Neisseria meningitidis complete ORF1 sequence.
 DE Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW Neisseria meningitidis; Neisseria infection; meningitis; septicemia; gonorrhea; ss.
 XX Neisseria meningitidis.
 OS Neisseria meningitidis.
 PN W03924578-A2.
 XX
 XX 20-MAY-1999.
 XX
 XX 03-OCT-1998; 98MO-1B001665.
 XX
 XX 06-NOV-1997; 97GB-00023516.
 XX 14-NOV-1997; 97GB-00024190.
 XX 18-NOV-1997; 97GB-00024386.
 XX 27-NOV-1997; 97GB-00025158.
 XX 10-DEC-1997; 97GB-00026147.
 XX 14-JAN-1998; 98GB-00000759.
 XX 01-SEP-1998; 98GB-00019016.
 XX
 XX (CHIR-) CHIRON SPA.
 XX
 XX Masignani V, Rapuoli R, Pizza M, Scarlato V, Grandi G;
 PI WPI: 1999-327407/27.
 DR P-PSDB; AAY38823.
 DR
 PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
 PT diagnosis, treatment and prevention of infection.
 XX
 PS Claim 9; Page 360-361; 524pp; English.
 XX
 XX Nucleotide sequences A211972-212358 represent open reading frames (ORFs)
 CC of Neisseria meningitidis and N. gonorrhoeae which encode antigenic
 CC proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
 CC their nucleic acids and antibodies are used for diagnosis, prevention (as
 CC vaccines) or treatment of Neisseria infections, such as meningitis,
 CC septicemia and gonorrhea. Both organisms are closely related. Fragments
 CC of the nucleic acids are useful as hybridisation probes and antisense
 CC reagents
 CC
 SC Sequence 4374 BP; 1266 A; 1163 C; 1066 G; 873 T; 0 U; 0 Other;
 Query Match 28.9%; Score 1245.4; DB 2; Length 4374;
 Best Local Similarity 58.8%; Pred. No. 3.2e-291;
 Matches 2565; Conservative 0; Mismatches 1536; Indels 264; Gaps 13;

QY 441 AACTACAAAAAG---ATATATTACCTCTTATGAGCGATTACCATACACGATTA 497
 DB 433 AATTATTAAGAGGACTTAAGGCCATCTTATGAGCGGATTAATATGCGCGTTTG 492
 QY 498 CATTAATTCCTTACAGAGCGGCTCCATTTGATATGATCTTGATATGATGCGACT 557
 DB 493 CATTAATTCCTTACAGAGCGGCTCCATTTGATATGATCTTGATATGATGCGACT 552
 QY 558 TATTCAGATGAGCAAAATATCCAGAAAGTTCGATCGGCTCGACCGGAGTCTTGG 617
 DB 553 TATTCAGATGAGCAAAATATCCAGAAAGTTCGATCGGCTCGACCGGAGTCTTGG 612
 QY 618 CGAATGATCAGACAAAGGCG-----GACCAAGTTCGCGTCAAT 659
 DB 613 CGATCTGATGAGATGAGCCCAATACCGGAAAGTTCATATCATTTGCAAGTGGAT 672
 QY 660 CATTAATTCCTTACAGAGCGGCTCCATTTGATATGATCTTGATATGATGCGACT 719
 DB 673 TCTTGGCTGCTGTTGCTGCGCATACCTTTGACAAATATGATGATGCTGCGACATCAAC 732
 QY 720 TTGGAGGCGGATGCTGTTAAAGCGGAGAAATATGCTCATACCGATTGCGAGGCTCAAG 779
 DB 733 TTAGGTATGAGAAATATTAACATAGCCCATATGCTTTTACCAACAGAGGCTCATTT 792
 QY 780 GGGACAGTGGTCTCCGATGTTATTTATGATGCTGAAAACAAAATGTTAAT 839
 DB 793 GGGACAGTGGCTCACCAATGTTATCTATGATGCGCCAAAGAAAGTGTATTAAT 852
 QY 840 GGGATATTCGGAGAGGCAACCTTTTGAAGCAAGAAATGCGTTCAATTTGTTGTC 899
 DB 853 GGGATATTCGGAGAGGCAACCTTTTGAAGCAAGAAATGCGTTCAATTTGTTGTC 912
 QY 900 AATCTTATTT---TGATGAAATTTTGGAAAGATTTATCATATCATCTTACACCGA 956
 DB 913 AAGATATGTTCTATATGAAATCTTTCGAGATACCATCATGATCTACAGACA 972
 QY 957 GCTGTAAATGAGATGACAAATATGAGAAATATGATGCTCAAGGCTTATATCTGAG 1016
 DB 973 CGTCAAAATGAGAAATATCTTTTACAGAGATATATATGAGCAAGAAATATCAATGCC 1032
 QY 1017 AATATGAG-----ATACATCGAATTAATTAATGATGATGAGTAT 1058
 DB 1033 AATATGAG-----ATACATCGAATTAATTAATGATGATGAGTAT 1092
 QY 1059 ATGATTAATCTTGAAGAGAAAGATTAATGATGATGATGATGATGATGATGATGAT 1118
 DB 1093 GTTCTTATTCGAG 1145
 QY 1119 ATTATTCCTCAGGTTTAAACAAATGAGAAAGCTATATTTATGATGATGATGATGAT 1178
 DB 1150 AGTTATCAGACGAGACTGATATATGAGAAATATTTCTTTATGAGAGAGAAAGGC 1209
 QY 1179 TCATTAATTCCTCAGATGATGATTAACAAAGGCGGAGTCTTATTTGAGGAGTAT 1238
 DB 1210 GAATGATATCTTACAGCAACATCATATATGATGATGATGATGATGATGATGATGAT 1269
 QY 1239 TTTACAGATCTTCCAAATTTCTAACAAATTTGCAAGAGGCTGATATCATGATGATGAT 1298
 DB 1270 TTTACAGATCTTCCAAATTTCTAACAAATTTGCAAGAGGCTGATATCATGATGATGAT 1329
 QY 1299 AATGACCGGTTACTTGGAAAGTAAATGAGCTGAGACATGATGATGATGATGATGATGAT 1358
 DB 1330 GACGATACCGGTTACTTGGAAAGTAAATGAGCTGAGACATGATGATGATGATGATGATGAT 1389
 QY 1359 AATGACCGGTTACTTGGAAAGTAAATGAGCTGAGACATGATGATGATGATGATGATGAT 1418
 DB 1390 AATGACCGGTTACTTGGAAAGTAAATGAGCTGAGACATGATGATGATGATGATGATGATGAT 1449
 QY 1419 GGTAAAGTATTTTGAAG 1478
 DB 1450 GGTAAAGTATTTTGAAG 1509
 QY 1479 ATTGGCTGTTTGAAG 1538

Qy	3540	GATGCGTCGGTCTATATGAGAGCAAAAAGCACTTAAGTAAATTGGGGTGCAAAA	3599
Db	3673	CAGATTTCGGCGCTACCGCCAC--AAACGACCTGGCCAAATGGATATCGAANA	3729
Qy	3600	GCCATTAGCTAATGAGCGAATTG993CAGTTTTCTCGATAGACGCTTCAGATTAACTTT	3659
Db	3730	AACTTCGGACGGGGCGCTGGCATCTCTTTTTCGACAAACGACGAAACAACCTTC	3789
Qy	3660	GATGACAGGTAAAAATACGCGACATTAAGATATGTCGGTTTTGGCCAAATTCAA	3719
Db	3790	GAGACGGCATGGCAACTCGCGACGCTTGCCACGCGCGCTTTTG993CAATTCGCG	3849
Qy	3720	TGGGGCAATTACATTTTGGTGTAAAGTGGAGCGGGAATCAGTCGACGTAAATGCGT	3779
Db	3850	ATGACAGGTTCTACATCGGCATACAGCGCGGGGCGGTTTTAGACAGGGACGCTTTCA	3909
Qy	3780	GAGAACAAACCGCAAAAATTCATCGAAAAGCATTAATTATGGCGTAAATGCAGTTT	3839
Db	3910	GAGCGCATGGAGGCAAAAATCCGCGCGCGGTGCGATTACGGCAATTCAGACAGATTC	3969
Qy	3840	CAGTTCGGTTAAG393CAATTG993CAATTCAGCCTTAATTTGAGATTAATCGATTTAT	3889
Db	3970	CGGCGCGGTTTCGGCGGATTCGGCATGGAACGCGCAATCG993CAACGGCTATTGTC	4029
Qy	3900	GAAAGTAAATTAATCAATTGAGGAATGAGAGTAAACGCTTGCTTGCATTAT	3959
Db	4030	CAAAAAGCGATTCGGCTACCAAAAAGTAAATATGCAACCCCGCGCTTCGATTCAAC	4089
Qy	3960	CGCTAAATCGCGCATGAGTTGATTAATTAATTAATTCGACAGATTAATACAGST	4019
Db	4090	CGCTACCGCGGGCGATTAAGGCAATTAATTCATTCACAAACCGCGCAACATTTGCAC	4149
Qy	4020	AAGCCTATTTCTTGTCATATTGTTGATGATTTTAAACGCTAACGTACAAACGAGTA	4079
Db	4150	ACGCTTATTGACCTGCTCATACGATGCGCGCTTG993CAAAATTCGCAACGCGTC	4209
Qy	4080	AATTCACGCGTTGCAACAACTTGGACGTTATTTGGCAAAAAGAATGGGATTAAG	4139
Db	4210	AATACCGCGTATTTGCTCAGATTTG993CAAAACCGCAGTGGGATTTGGGCGTAAAC	4269
Qy	4140	GCGAATATTTACATTTCCAAATTTCCGCTTTATCTCAAAATCTCAAGTTTCACAATC	4199
Db	4270	GCGAATCAAAAGGTTTCAAGCTGTCCTCCACAGCTGCGCGCAAAAGCCCGCACTG	4329
Qy	4200	GCGAAACAGCAAAATGTCGGCGTGAATTTGGGTATAGTTGGTAA	4244
Db	4330	GAGCGCAACAGCGCGGCGATCAAAATTGGGCTACGCGCTGTA	4374

Search completed: June 24, 2004, 08:41:45
Job time : 1593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 13:12:32 ; Search time 1688 Seconds

(without alignments)
11721.401 Million cell updates/sec

Title: US-10-645-655-1

Perfect score: 4319

Sequence: 1 TCAATAGTCGTTACTAGT.....ACCTTTATCCCATAT 4319

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/FCUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4319	100.0	4319	US-09-839-996-1	Sequence 1, Appli
2	4319	100.0	4319	US-10-645-655-1	Sequence 1, Appli
3	4319	100.0	4319	US-10-080-505-1	Sequence 1, Appli
4	3180.8	73.6	4305	US-10-080-505-8	Sequence 8, Appli
5	3019.8	69.9	1830121	US-10-329-960-1	Sequence 1, Appli
6	3019.8	69.9	1830121	US-10-329-960-1	Sequence 1, Appli
7	2990.4	69.2	5245	US-10-080-505-12	Sequence 12, Appli
8	2970.2	64.1	4828	US-10-080-505-16	Sequence 16, Appli
9	2770.2	64.1	4828	US-10-080-505-14	Sequence 14, Appli
10	2709.6	62.7	4350	US-10-080-505-10	Sequence 10, Appli
11	1353.8	32.3	4350	US-10-282-122A-29527	Sequence 29527, A
12	1294.8	30.0	4404	US-10-282-122A-28843	Sequence 28843, A
13	202	4.7	5085	US-10-282-122A-22193	Sequence 22193, A
14	164.4	3.8	5322	US-10-282-122A-29633	Sequence 29633, A

15	155.2	3.6	4779	13	US-10-282-122A-29078	Sequence 29078, A
16	126.8	2.9	4113	15 <th>US-10-238-075-890</th> <th>Sequence 829, App</th>	US-10-238-075-890	Sequence 829, App
17	126.8	2.9	48012	15 <th>US-10-C85-953-251</th> <th>Sequence 251, App</th>	US-10-C85-953-251	Sequence 251, App
18	126.8	2.9	48254	15 <th>US-10-238-075-829</th> <th>Sequence 829, App</th>	US-10-238-075-829	Sequence 829, App
19	100.6	2.3	4128	15 <th>US-10-238-075-559</th> <th>Sequence 559, App</th>	US-10-238-075-559	Sequence 559, App
20	100.6	2.3	7654	15 <th>US-10-238-075-554</th> <th>Sequence 554, App</th>	US-10-238-075-554	Sequence 554, App
21	100.6	2.3	7654	15 <th>US-10-C85-953-15</th> <th>Sequence 15, Appl</th>	US-10-C85-953-15	Sequence 15, Appl
22	45.8	1.1	6378	15 <th>US-10-311-455-149</th> <th>Sequence 149, App</th>	US-10-311-455-149	Sequence 149, App
23	45.4	1.1	6283	15 <th>US-10-311-455-61</th> <th>Sequence 61, Appl</th>	US-10-311-455-61	Sequence 61, Appl
24	44.6	1.0	564	13 <th>US-10-027-632-50664</th> <th>Sequence 50664, A</th>	US-10-027-632-50664	Sequence 50664, A
25	44.6	1.0	564	13 <th>US-10-027-632-50665</th> <th>Sequence 50665, A</th>	US-10-027-632-50665	Sequence 50665, A
26	44.6	1.0	564	13 <th>US-10-027-632-66877</th> <th>Sequence 66877, A</th>	US-10-027-632-66877	Sequence 66877, A
27	44.6	1.0	564	13 <th>US-10-027-632-66878</th> <th>Sequence 66878, A</th>	US-10-027-632-66878	Sequence 66878, A
28	44.6	1.0	564	13 <th>US-10-027-632-68531</th> <th>Sequence 68531, A</th>	US-10-027-632-68531	Sequence 68531, A
29	44.6	1.0	564	13 <th>US-10-027-632-68532</th> <th>Sequence 68532, A</th>	US-10-027-632-68532	Sequence 68532, A
30	44.6	1.0	564	13 <th>US-10-027-632-70179</th> <th>Sequence 70179, A</th>	US-10-027-632-70179	Sequence 70179, A
31	44.6	1.0	564	13 <th>US-10-027-632-70180</th> <th>Sequence 70180, A</th>	US-10-027-632-70180	Sequence 70180, A
32	44.6	1.0	564	13 <th>US-10-027-632-296122</th> <th>Sequence 296122, A</th>	US-10-027-632-296122	Sequence 296122, A
33	44.6	1.0	564	13 <th>US-10-027-632-296123</th> <th>Sequence 296123, A</th>	US-10-027-632-296123	Sequence 296123, A
34	44.6	1.0	564	16 <th>US-10-027-632-50664</th> <th>Sequence 50664, A</th>	US-10-027-632-50664	Sequence 50664, A
35	44.6	1.0	564	16 <th>US-10-027-632-50665</th> <th>Sequence 50665, A</th>	US-10-027-632-50665	Sequence 50665, A
36	44.6	1.0	564	16 <th>US-10-027-632-66877</th> <th>Sequence 66877, A</th>	US-10-027-632-66877	Sequence 66877, A
37	44.6	1.0	564	16 <th>US-10-027-632-66878</th> <th>Sequence 66878, A</th>	US-10-027-632-66878	Sequence 66878, A
38	44.6	1.0	564	16 <th>US-10-027-632-68531</th> <th>Sequence 68531, A</th>	US-10-027-632-68531	Sequence 68531, A
39	44.6	1.0	564	16 <th>US-10-027-632-68532</th> <th>Sequence 68532, A</th>	US-10-027-632-68532	Sequence 68532, A
40	44.6	1.0	564	16 <th>US-10-027-632-70179</th> <th>Sequence 70179, A</th>	US-10-027-632-70179	Sequence 70179, A
41	44.6	1.0	564	16 <th>US-10-027-632-70180</th> <th>Sequence 70180, A</th>	US-10-027-632-70180	Sequence 70180, A
42	44.6	1.0	564	16 <th>US-10-027-632-296122</th> <th>Sequence 296122, A</th>	US-10-027-632-296122	Sequence 296122, A
43	44.6	1.0	564	16 <th>US-10-027-632-296123</th> <th>Sequence 296123, A</th>	US-10-027-632-296123	Sequence 296123, A
44	44.6	1.0	654	13 <th>US-10-027-632-276009</th> <th>Sequence 276009, A</th>	US-10-027-632-276009	Sequence 276009, A
45	44.6	1.0	654	16 <th>US-10-027-632-276009</th> <th>Sequence 276009, A</th>	US-10-027-632-276009	Sequence 276009, A

ALIGNMENTS

RESULT 1
US-09-839-996-1
Sequence 1, Application US/09839996
Publication No. US20030009010A1
GENERAL INFORMATION:
APPLICANT: St. Gene III, Joseph W.
TITLE OF INVENTION: Hemophilus Adherence and Penetration Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Eliehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,996
FILING DATE: 20-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RFT/RYS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1899
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4319 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: both

FEATURE:

NAME/KEY: CDS

LOCATION: 60..4241

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-B39-996-1

Query Match 100.0%; Score 4319; DB 10; Length 4319;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCATATGCTTTAATCTAGATTTTAAATGAGAAATTTCTTAATTAATTAACCTTA
 1 TCATATGCTTTAATCTAGATTTTAAATGAGAAATTTCTTAATTAATTAACCTTA
 61 TGAATAAACTGATTTGCTCTTAATTTTAAACGCTTGATTTCAATGAGATAT 120
 61 TGAATAAACTGATTTGCTCTTAATTTTAAACGCTTGATTTCAATGAGATAT 120
 121 CGCAAGCGTGGCTGCTCACTTAATTTTGGATTTGATTACCAATTAATCTGATTTG 180
 121 CGCAAGCGTGGCTGCTCACTTAATTTTGGATTTGATTACCAATTAATCTGATTTG 180
 181 CGGAGATTAAGGAAAGTTCAAGTTGGGCTCAAAATTAATTAAGCTTTAATCAACAG 240
 181 CGGAGATTAAGGAAAGTTCAAGTTGGGCTCAAAATTAATTAAGCTTTAATCAACAG 240
 241 GCGAATTAATGCGAATCAATTAAGCAAAAGCCCATTAATTAATTAATTAATTAAT 300
 241 GCGAATTAATGCGAATCAATTAAGCAAAAGCCCATTAATTAATTAATTAATTAAT 300
 301 GAAACGCGTGGCAGCTTGTGTAATTAATTAATTAATTAATTAATTAATTAATTA 360
 301 GAAACGCGTGGCAGCTTGTGTAATTAATTAATTAATTAATTAATTAATTAATTA 360
 361 GATATACAGATTTGATTTTGGTGCAGAGGAAACCAACCCGATCAATCTGTTACTT 420
 361 GATATACAGATTTGATTTTGGTGCAGAGGAAACCAACCCGATCAATCTGTTACTT 420
 421 AATAGATTTGAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 421 AATAGATTTGAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 481 ACCATTAATCAAGATTAATTAATTTGTTACAGAGCGCTCCAAATTAATTAATTAAT 540
 481 ACCATTAATCAAGATTAATTAATTTGTTACAGAGCGCTCCAAATTAATTAATTAAT 540
 541 AATAGATTTGAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 541 AATAGATTTGAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 601 CTGAGCGCAGTTTGGGAAATGATCAAGCAAAAGCAAGCAATGTCGCGGTCATTC 660
 601 CTGAGCGCAGTTTGGGAAATGATCAAGCAAAAGCAAGCAATGTCGCGGTCATTC 660
 661 ATTATCTGACAGCTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
 661 ATTATCTGACAGCTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
 721 TGGAGGCGATTTGCTTAAGCGGAAATTAATTAATTAATTAATTAATTAATTAATTA 780
 721 TGGAGGCGATTTGCTTAAGCGGAAATTAATTAATTAATTAATTAATTAATTAATTA 780
 781 GGGACAGTGGTTCCGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
 781 GGGACAGTGGTTCCGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
 841 GGAATATTACGGGAAAGCAACCTTTTGAAGCAAGAAATTAATTAATTAATTAATTA 900

841 GGAATATTACGGGAAAGCAACCTTTTGAAGCAAGAAATTAATTAATTAATTAATTA 900
 901 AATCTATTATTTGAAATTTTGAAGATTAATTAATTAATTAATTAATTAATTAATTA 960
 901 AATCTATTATTTGAAATTTTGAAGATTAATTAATTAATTAATTAATTAATTAATTA 960
 961 GATATGAGTGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
 961 GATATGAGTGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
 1021 CAGAAATACATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
 1021 CAGAAATACATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
 1081 AGATTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
 1081 AGATTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
 1141 ATGAGAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
 1141 ATGAGAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
 1201 TTAACCAAGGCGGCTGCTTAATTTTGAAGGATTAATTAATTAATTAATTAATTA 1260
 1201 TTAACCAAGGCGGCTGCTTAATTTTGAAGGATTAATTAATTAATTAATTAATTA 1260
 1261 ACCAATCTTGGCAGAGCTGGGATTAATTAATTAATTAATTAATTAATTAATTAAT 1320
 1261 ACCAATCTTGGCAGAGCTGGGATTAATTAATTAATTAATTAATTAATTAATTAAT 1320
 1321 TAATGAGTGGAAACATGATCACTTTCTAAATTTGGTAAAGAAATTAATTAATTAAT 1380
 1321 TAATGAGTGGAAACATGATCACTTTCTAAATTTGGTAAAGAAATTAATTAATTAAT 1380
 1381 CCAAGGCGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1440
 1381 CCAAGGCGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1440
 1441 AGGCAAGATTAAGGCAACCAAGCTTTAGTAAATTTGGCTTGGCTGAGAG 1500
 1441 AGGCAAGATTAAGGCAACCAAGCTTTAGTAAATTTGGCTTGGCTGAGAG 1500
 1501 GGAATCTTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
 1501 GGAATCTTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
 1561 GTGATGCTGCTGATTTAAGGCAATTAATTAATTAATTAATTAATTAATTAATTA 1620
 1561 GTGATGCTGCTGATTTAAGGCAATTAATTAATTAATTAATTAATTAATTAATTA 1620
 1621 ACCGAGGCGCAATGATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
 1621 ACCGAGGCGCAATGATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
 1681 GGAACGAAGCATTTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1740
 1681 GGAACGAAGCATTTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1740
 1741 AATATGCTTAACAGCTTTGGGCAATTAATTAATTAATTAATTAATTAATTAATTA 1800
 1741 AATATGCTTAACAGCTTTGGGCAATTAATTAATTAATTAATTAATTAATTAATTA 1800
 1801 AATATGCTTAACAGCTTTGGGCAATTAATTAATTAATTAATTAATTAATTAATTA 1860
 1801 AATATGCTTAACAGCTTTGGGCAATTAATTAATTAATTAATTAATTAATTAATTA 1860
 1861 TAAAGGCGATTTACCAAAACAAAGGTAATTTTCAAGGCTGAGACGACCG 1920
 1861 TAAAGGCGATTTACCAAAACAAAGGTAATTTTCAAGGCTGAGACGACCG 1920
 1921 ACGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1980

Db 1921 AGCCTACATCATTTAATAAGTGGTCAAGAAATGGAAGTATACAAAGGCGAA 1980
Qy 1981 TTGTGTGGATGCAATGGATGCAACGGTACATTTAAAGTGAACCTTCGAAATTAAG 2640
Db 1981 TTGTGTGGATGCAATGGATGCAACGGTACATTTAAAGTGAACCTTCGAAATTAAG 2040
Qy 2041 GGGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2100
Db 2041 GGGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2100
Qy 2101 ATATGCAAAATGCAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2160
Db 2101 ATATGCAAAATGCAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2160
Qy 2161 CAGATTGCAAGATTAACGATTTGCAAAAAGTGAATTAACGATTAACGATTTGCA 2220
Db 2161 CAGATTGCAAGATTAACGATTTGCAAAAAGTGAATTAACGATTAACGATTTGCA 2220
Qy 2221 ATTCTATTCGAAATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
Db 2221 ATTCTATTCGAAATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
Qy 2281 ATGTTAAGTTAGCAAAATTAATGCAATGCAATTAATGCAATTAATGCAATTAATGCA 2340
Db 2281 ATGTTAAGTTAGCAAAATTAATGCAATGCAATTAATGCAATTAATGCAATTAATGCA 2340
Qy 2341 CATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2400
Db 2341 CATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2400
Qy 2401 CGGTGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2460
Db 2401 CGGTGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2460
Qy 2461 TAAAAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2520
Db 2461 TAAAAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2520
Qy 2521 AAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2580
Db 2521 AAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2580
Qy 2581 GTCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2640
Db 2581 GTCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2640
Qy 2641 CATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2700
Db 2641 CATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2700
Qy 2701 ATGTTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2760
Db 2701 ATGTTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2760
Qy 2761 GCGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2820
Db 2761 GCGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2820
Qy 2821 CAGGCAAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2880
Db 2821 CAGGCAAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2880
Qy 2881 CGTTATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2940
Db 2881 CGTTATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2940
Qy 2941 GTTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3000
Db 2941 GTTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3000
Qy 3001 AATTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3060
Db 3001 AATTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3060

Qy 3061 TTGAACGACTGCTTAAACCAAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3120
Db 3061 TTGAACGACTGCTTAAACCAAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3120
Qy 3121 GAGGAGGTTTCTGATACCTGCTGATCAAGCTGTTAAACGATTAAGCAAC 3180
Db 3121 GAGGAGGTTTCTGATACCTGCTGATCAAGCTGTTAAACGATTAAGCAAC 3180
Qy 3181 AAGTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3240
Db 3181 AAGTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3240
Qy 3241 AAAGGAGGTTTCTGATACCTGCTGATCAAGCTGTTAAACGATTAAGCAAC 3300
Db 3241 AAAGGAGGTTTCTGATACCTGCTGATCAAGCTGTTAAACGATTAAGCAAC 3300
Qy 3301 TTGAGTTATGATGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3360
Db 3301 TTGAGTTATGATGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3360
Qy 3361 AAAAGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3420
Db 3361 AAAAGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3420
Qy 3421 TATCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3480
Db 3421 TATCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3480
Qy 3481 ATCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3540
Db 3481 ATCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3540
Qy 3541 ATGCTG 3600
Db 3541 ATGCTG 3600
Qy 3601 CTTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3660
Db 3601 CTTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3660
Qy 3661 ATGAACGTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3720
Db 3661 ATGAACGTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3720
Qy 3721 GGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3780
Db 3721 GGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3780
Qy 3781 AAGAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3840
Db 3781 AAGAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3840
Qy 3841 AGTTCCGTTTAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3900
Db 3841 AGTTCCGTTTAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3900
Qy 3901 AAGTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3960
Db 3901 AAGTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3960
Qy 3961 GCTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 4020
Db 3961 GCTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 4020
Qy 4021 AGCTTATTTCTGCTGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 4080
Db 4021 AGCTTATTTCTGCTGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 4080
Qy 4081 ATCTACGCTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 4140
Db 4081 ATCTACGCTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 4140


```

Db      3361 AAAAACAAGCAAAAGAAAGCTTGAATGACCGTATTCAAAATAGTGTATCGAAT 3420
Qy      3421 TATTCGAAACAGTAATATGATGCTTTCTGTTCAAGATGAATGATGCTTTTGTAG 3480
      3421 TATTCGAAACAGTAATATGATGCTTTCTGTTCAAGATGAATGATGCTTTTGTAG 3480
Qy      3481 ATGACCAACATGCGCGGTGAGCAAAATATGCAAGGATTAAGAGAGCGTATGATC 3540
      3481 ATGACCAACATGCGCGGTGAGCAAAATATGCAAGGATTAAGAGAGCGTATGATC 3540
Qy      3541 ATGCGTTCGCTTATGACAGCAGAGAAACGAACTTACGTAATGCGGTGCAAAAG 3600
      3541 ATGCGTTCGCTTATGACAGCAGAGAAACGAACTTACGTAATGCGGTGCAAAAG 3600
Qy      3601 CTTAGCTATGAGCAATGCGGCACTTTCTGCGATGACGCTTCAATATACCTT 3660
      3601 CTTAGCTATGAGCAATGCGGCACTTTCTGCGATGACGCTTCAATATACCTT 3660
Qy      3661 ATGACCAAGTAAATATGACGCAATTAAGATGATGATGATGATGATGATGAT 3720
      3661 ATGACCAAGTAAATATGACGCAATTAAGATGATGATGATGATGATGATGAT 3720
Qy      3721 GGGCGCATTTACATTTGTTGTAACGTGGAACGGGATCAGTGGAGTAAATGCTG 3780
      3721 GGGCGCATTTACATTTGTTGTAACGTGGAACGGGATCAGTGGAGTAAATGCTG 3780
Qy      3781 AAGAACAAAGCGGAAATATGATGGAAGGATTAATATGCGGTGATGATGATGAT 3840
      3781 AAGAACAAAGCGGAAATATGATGGAAGGATTAATATGCGGTGATGATGATGAT 3840
Qy      3841 AGTTCGTTAGGCGCATTTGCGCATTTGCGCATTTGCGCATTTGCGCATTTGCG 3900
      3841 AGTTCGTTAGGCGCATTTGCGCATTTGCGCATTTGCGCATTTGCGCATTTGCG 3900
Qy      3901 AACGTAAATTTATCAATCTGAGGAGTGAAGTGAAGGAGGCTTACATTAATC 3960
      3901 AACGTAAATTTATCAATCTGAGGAGTGAAGTGAAGGAGGCTTACATTAATC 3960
Qy      3961 GGTATATGCTGCAATGATGATTAATTAATTAATTAATTAATTAATTAATTA 4020
      3961 GGTATATGCTGCAATGATGATTAATTAATTAATTAATTAATTAATTAATTA 4020
Qy      4021 AACCTATTTCTGCAATGATGATTAATTAATTAATTAATTAATTAATTAAT 4080
      4021 AACCTATTTCTGCAATGATGATTAATTAATTAATTAATTAATTAATTAAT 4080
Qy      4081 ATCTACGCGTGTGCAACCACTTGAACGTTATGCGAAGGATTAAGG 4140
      4081 ATCTACGCGTGTGCAACCACTTGAACGTTATGCGAAGGATTAAGG 4140
Qy      4141 CAGAAATTTATCAATTTCCCAATTTCCGCTTTATCTCAAAATCTCAAGTTCA 4200
      4141 CAGAAATTTATCAATTTCCCAATTTCCGCTTTATCTCAAAATCTCAAGTTCA 4200
Qy      4201 GGAACAGCAAAATGCGCGTGAATTTGCGATGATGATGATGATGATGATGAT 4260
      4201 GGAACAGCAAAATGCGCGTGAATTTGCGATGATGATGATGATGATGATGAT 4260
Qy      4261 ATGCTTTATGATTAACAAGTGTGATGATGATGATGATGATGATGATGATGAT 4319
      4261 ATGCTTTATGATTAACAAGTGTGATGATGATGATGATGATGATGATGATGAT 4319
      4261 ATGCTTTATGATTAACAAGTGTGATGATGATGATGATGATGATGATGATGAT 4319

```

```

; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (60)..(4241)
; OTHER INFORMATION:
US-10-380-505-1

Query Match      100.0%; Score 4319; DB 15; Length 4319;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TCATATGCTTTTACTAGTATTTTATACGAAATTAATTAATTAATTAATTAATTA 60
Db      1 TCATATGCTTTTACTAGTATTTTATACGAAATTAATTAATTAATTAATTAATTA 60
Qy      61 TGAATAAAACGTAATTTGCTTAAATTTTAAACGCTGCAATTCATTAAGGATGAT 120
Db      61 TGAATAAAACGTAATTTGCTTAAATTTTAAACGCTGCAATTCATTAAGGATGAT 120
Qy      121 CGAAGCTGGGCTGTGCACTTATTTTGGATGATGATGATGATGATGATGATGAT 180
Db      121 CGAAGCTGGGCTGTGCACTTATTTTGGATGATGATGATGATGATGATGATGAT 180
Qy      181 CGAGATTAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db      181 CGAGATTAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Qy      241 GGCATATGTTGAGCACTCAATGACAAAGCCCGATGATGATGATGATGATGATGAT 300
Db      241 GGCATATGTTGAGCACTCAATGACAAAGCCCGATGATGATGATGATGATGATGAT 300
Qy      301 GTACGCGGTGGAGCTTGTGTAATTAATTAATTAATTAATTAATTAATTAATTA 360
Db      301 GTACGCGGTGGAGCTTGTGTAATTAATTAATTAATTAATTAATTAATTAATTA 360
Qy      361 GATATACAGATGATTTTGTGCAAGGAAACCCCGATGATGATGATGATGATGAT 420
Db      361 GATATACAGATGATTTTGTGCAAGGAAACCCCGATGATGATGATGATGATGAT 420
Qy      421 ATAAATGTTAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
Db      421 ATAAATGTTAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
Qy      481 ACCATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
Db      481 ACCATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
Qy      541 AATGATGCGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Db      541 AATGATGCGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Qy      601 CTGAGCGGAGTTTGGCGAATGATCAAGACAAAGCGCAAGTGGCGGTGATGAT 660
Db      601 CTGAGCGGAGTTTGGCGAATGATCAAGACAAAGCGCAAGTGGCGGTGATGAT 660
Qy      661 AATATGACAGCTGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Db      661 AATATGACAGCTGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Qy      721 TGGAGGCGATGTTGTAACCGGAGAAATGATGATGATGATGATGATGATGATGAT 780
Db      721 TGGAGGCGATGTTGTAACCGGAGAAATGATGATGATGATGATGATGATGATGAT 780
Qy      781 GGGACAGTGTCTCGAGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 840

```

Db 781 GGGACAGTGGTTCGCCAGTGTATTTATGATGCTGAAAAAAGTGTATTAAG 840
Qy 841 GGAATTAAGGGAAGGCAACCTTTGAGGCAAGAAAAAGGTTCAATGGTTCGA 900
Db 841 GGAATTAAGGGAAGGCAACCTTTGAGGCAAGAAAAAGGTTCAATGGTTCGA 900
Qy 901 AATCTTATTTGATGAAATTTTCGAAAGATTTATACATCACTTACACCCGAGCTG 960
Db 901 AATCTTATTTGATGAAATTTTCGAAAGATTTATACATCACTTACACCCGAGCTG 960
Qy 961 GTAATGAGTGTACACATTTAGTGAATGTAATGTGTAGGGCTTATTAACGAAAT 1020
Db 961 GTAATGAGTGTACACATTTAGTGAATGTAATGTGTAGGGCTTATTAACGAAAT 1020
Qy 1021 CAGGAATACCTACGAAATTTAAATTTAGCTGAAATATGAGTTTACCTTTGAAAGAGA 1080
Db 1021 CAGGAATACCTACGAAATTTAAATTTAGCTGAAATATGAGTTTACCTTTGAAAGAGA 1080
Qy 1081 AGGATTAAGTTCATATCTAGATATGACGCACTTAATTTATCTTCCAGTTTAAACA 1140
Db 1081 AGGATTAAGTTCATATCTAGATATGACGCACTTAATTTATCTTCCAGTTTAAACA 1140
Qy 1141 ATGAGAAACCTTATATTTATGATCAAAAAAGATCATTAATCTTGGCATCGACA 1200
Db 1141 ATGAGAAACCTTATATTTATGATCAAAAAAGATCATTAATCTTGGCATCGACA 1200
Qy 1201 TTACCAAGGGGCGGGGTCTTATTTGAGGTAATTTTACAGTATCTCCAAATTTCTA 1260
Db 1201 TTACCAAGGGGCGGGGTCTTATTTGAGGTAATTTTACAGTATCTCCAAATTTCTA 1260
Qy 1261 ACCAACTTGGCAGAGGCTGGCATATGTAATGTAATTAAGCAACGTTACTTGGAAAG 1320
Db 1261 ACCAACTTGGCAGAGGCTGGCATATGTAATGTAATTAAGCAACGTTACTTGGAAAG 1320
Qy 1321 TAAATGCGTGGAAATGATGATCTTCTAAATTTGTAAGAAATTTGACGTTTCAAG 1380
Db 1321 TAAATGCGTGGAAATGATGATCTTCTAAATTTGTAAGAAATTTGACGTTTCAAG 1380
Qy 1381 CCAAAGGGGAAAAATTAAGTGTGATGATGAGGCAATGTAATGTAATTTGTAAGCAAG 1440
Db 1381 CCAAAGGGGAAAAATTAAGTGTGATGATGAGGCAATGTAATGTAATTTGTAAGCAAG 1440
Qy 1441 AGGCAAGCATCAAGGCAACAAAGCCTTATGTAATTTGCTTGGTTCAGCGCAAG 1500
Db 1441 AGGCAAGCATCAAGGCAACAAAGCCTTATGTAATTTGCTTGGTTCAGCGCAAG 1500
Qy 1501 GGAATGCTGATTAAGATGATGATTAACATTTGATACCGATTAATTTTCCGCTTTC 1560
Db 1501 GGAATGCTGATTAAGATGATGATTAACATTTGATACCGATTAATTTTCCGCTTTC 1560
Qy 1561 GTGGTGTGCTTATGATCTTAAAGGCAATCACTTAACCTTAAAGGTAATCCAAATATACG 1620
Db 1561 GTGGTGTGCTTATGATCTTAAAGGCAATCACTTAAAGGTAATCCAAATATACG 1620
Qy 1621 AGGAGGGGCAATGATTTGTGAACCAATTAACATCAAGCCGTATGTCATTAATCTG 1680
Db 1621 AGGAGGGGCAATGATTTGTGAACCAATTAACATCAAGCCGTATGTCATTAATCTG 1680
Qy 1681 GGAAGCAAGCATTTGTTCTACCTATGAAATATATTTAAATCTGATTAACGAAAG 1740
Db 1681 GGAAGCAAGCATTTGTTCTACCTATGAAATATATTTAAATCTGATTAACGAAAG 1740
Qy 1741 AAATGCTTACAGGTTGTTTGGGAAACAGATTAATTAACCAATGAGGATTA 1800
Db 1741 AAATGCTTACAGGTTGTTTGGGAAACAGATTAATTAACCAATGAGGATTA 1800
Qy 1801 ACCTTATTTATTAACCAACAGAGATGATCTTGTCTACTTTCAGTGGTCAAAAT 1860
Db 1801 ACCTTATTTATTAACCAACAGAGATGATCTTGTCTACTTTCAGTGGTCAAAAT 1860
Qy 1861 TAAAGGGATTTTACCAACCAAGGTAATTTTTCAGCGGTAGACCGCAACCCG 1920

Db 1861 TAAAGGGATTTTACCAACCAAGGTAATTTTTCAGCGGTAGACCGCAACCCG 1920
Qy 1921 AAGCTTCAATCATTTAAATTAACGTTTGTGAAATGAAAGTATACCAAGCCGAAA 1980
Db 1921 AAGCTTCAATCATTTAAATTAACGTTTGTGAAATGAAAGTATACCAAGCCGAAA 1980
Qy 1981 TTGTGTGGATCAGATTTGATCAACCGTACATTTAAAGCTGAAAACTTCCAAATTAAG 2040
Db 1981 TTGTGTGGATCAGATTTGATCAACCGTACATTTAAAGCTGAAAACTTCCAAATTAAG 2040
Qy 2041 GGGAGATGGGGTGTCTGCAATTTCTCAATTAAGGGAATTTGACAGTACGA 2100
Db 2041 GGGAGATGGGGTGTCTGCAATTTCTCAATTAAGGGAATTTGACAGTACGA 2100
Qy 2101 AATAAGCAATGCAATTTGTGTGTGCAATTAACCAATTAACATTTTCAAGCTT 2160
Db 2101 AATAAGCAATGCAATTTGTGTGTGCAATTTAACCAATTAACATTTTCAAGCTT 2160
Qy 2161 CAGATTTGACAGATTTACGATTTGCAAAAAAGTGAATTTACCGATTAACCAAGTATTA 2220
Db 2161 CAGATTTGACAGATTTACGATTTGCAAAAAAGTGAATTTACCGATTAACCAAGTATTA 2220
Qy 2221 ATTCTATACCAAAACCAAAATCAATGCTATTTATTTAATGATTAATGCAAGCGGA 2280
Db 2221 ATTCTATACCAAAACCAAAATCAATGCTATTTATTTAATGATTAATGCAAGCGGA 2280
Qy 2281 ATGTTAAAGTTTGAACAACTTAATGCAATGCACTTTAAACATCAACGCAATTTA 2340
Db 2281 ATGTTAAAGTTTGAACAACTTAATGCAATGCACTTTAAACATCAACGCAATTTA 2340
Qy 2341 CATTTAGCAACATGCAACCCAAATAGGCAATTTGCACTTCCGCAATTTCAACGCA 2400
Db 2341 CATTTAGCAACATGCAACCCAAATAGGCAATTTGCACTTCCGCAATTTCAACGCA 2400
Qy 2401 CGGTGGAATGCAACCTTGAACGTAATGCACTTTAAGGATTTAGCTCAATTTTCTT 2460
Db 2401 CGGTGGAATGCAACCTTGAACGTAATGCACTTTAAGGATTTAGCTCAATTTTCTT 2460
Qy 2461 TAAAAAACAGCAATTTTTCGCAACAAATTCAGGGAACAAAGGCAACAGTGCCTGG 2520
Db 2461 TAAAAAACAGCAATTTTTCGCAACAAATTCAGGGAACAAAGGCAACAGTGCCTGG 2520
Qy 2521 AAATGCACTTGGCAACATGCTTAGCATATCTGCAATTTAAGGTAATTAACCA 2580
Db 2521 AAATGCACTTGGCAACATGCTTAGCATATCTGCAATTTAAGGTAATTAACCA 2580
Qy 2581 GTAGCATCACTTAAATTCAGTTATTCAGTTATTCAGTTATTCAGTTATTCAGTTATTC 2640
Db 2581 GTAGCATCACTTAAATTCAGTTATTCAGTTATTCAGTTATTCAGTTATTCAGTTATTC 2640
Qy 2641 CATTAGAGAGGAAACAAAGCCCAACATGCGCAACATGCTTCAACCAATTAACGTA 2700
Db 2641 CATTAGAGAGGAAACAAAGCCCAACATGCGCAACATGCTTCAACCAATTAACGTA 2700
Qy 2701 ATGTAATTTAGTGGCAAGGCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCA 2760
Db 2701 ATGTAATTTAGTGGCAAGGCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCA 2760
Qy 2761 GCGATTAATTTAAATTTATTCATGACCTGAGGCAATTAATTTATCTGTTCGCAACA 2820
Db 2761 GCGATTAATTTAAATTTATTCATGACCTGAGGCAATTAATTTATCTGTTCGCAACA 2820
Qy 2821 CAGGCAAGAACCCGAAACCTTGTGCAATTTAATTTGTGGAAGCAAGATTAACAC 2880
Db 2821 CAGGCAAGAACCCGAAACCTTGTGCAATTTAATTTGTGGAAGCAAGATTAACAC 2880
Qy 2881 CGTATGAGATGATGCTCAATTTATTTAGAAATGACAGGATGATGAGGATTAAC 2940
Db 2881 CGTATGAGATGATGCTCAATTTATTTAGAAATGACAGGATGATGAGGATTAAC 2940
Qy 2941 GTTATTAATTTAGTGAAGATGATGAGGATTTCCGTTGATTAACCAATTAAGGCAAG 3000
Db 2941 GTTATTAATTTAGTGAAGATGATGAGGATTTCCGTTGATTAACCAATTAAGGCAAG 3000

Db 2515 TTTAAAAACATGCTATTTTTCGACCAATTCAGGCGCAAAAGACACAAAGTACGTTG 2574
 Qy 2520 GAAATGCGACTTGGACAAATGCTACCGATCTACATTCAGATTTAAAGCTAAATAC 2579
 Db 2575 GAAATGCGACTTGGACAAATGCTACCGATCTACATTCAGATTTAAAGCTAAATAC 2634
 Qy 2580 AGTACATCAGCTTAAATTCAGCTTATTCAGTACCTCAACAAATGCGCAGCT--CGC 2636
 Db 2635 AGTACATCAGCTTAAATTCAGCTTATTCAGTACCTCAACAAATGCGCAGCTCGC 2694
 Qy 2637 CGTTTCATGAGAGGAAACAAACGCAATCGCAGCAATCGTTTCAACATTCAGCA 2696
 Db 2695 CGTTTCATGAGAGGAAACAAACGCAATCGCAGCAATCGTTTCAACATTCAGCA 2754
 Qy 2697 GTAAATGTAATGAGTGGGCAAGGCAATTCATTCATTCCTTATTTGGCTAT 2756
 Db 2755 GTAAATGTAATGAGTGGGCAAGGCAATTCATTCATTCCTTATTTGGCTAT 2814
 Qy 2757 AAAAGCGATTAATTTAAATTTTCATATGCGTGAAGGCGCATTCATTTATCTTCCG 2816
 Db 2815 GAAAGCGATTAATTTAAATTTATCAATGACGCTGAGGCGCATTCATTTATCTTCCG 2874
 Qy 2817 AACACAGCAAGAACCCGAAACCTTGAAGCAATTAATCTTGGTTGAAAGCAAGATAT 2876
 Db 2875 AACACAGCAAGAACCCGAAACCTTGAAGCAATTAATCTTGGTTGAAAGCAAGATAT 2934
 Qy 2877 CAACCGTTACAGATTAAGCTCAAAATTTACTTTAGAAATGACCAAGTTGATGACGTTCA 2936
 Db 2935 CAACCGTTACAGATTAAGCTCAAAATTTACTTTAGAAATGACCAAGTTGATGACGTTCA 2994
 Qy 2937 TTACGTTTAAATTTAGTGAAGATGATGCGCAATTCGCTTGAATACCCAAATTAAGAG 2996
 Db 2995 TTACGTTTAAATTTAGTGAAGATGATGCGCAATTCGCTTGAATACCCAAATTAAGAG 3054
 Qy 2997 CAGGAATGCAATGATTTAGTGAAGAGCAGCAAGCAAGCAATTAAGAGCCAA 3056
 Db 3055 CAGGAATGCGCTCTGATTTAGTGAAGAGCAGCAAGCAAGCAATTAAGAGCCAA 3114
 Qy 3057 CAAGTTGAACCGACTGCTAAABCAAAACAGGTGAGCCAAAGTGGCTCAAGAGCA 3116
 Db 3115 CAAGTTGAACCGACTGCTAAABCAAAACAGGTGAGCCAAAGTGGCTCAAGAGCA 3168
 Qy 3117 GCGAGAGCAGCTTCTCGATACCTGCGCGATTCAAAGCTGTTAAAGCATTAAGAGCC 3176
 Db 3169 --AGACGGGTGTCTGATACCCGCTGCTCAAAAGCTGTTAAAGCATTAAGAGCC 3225
 Qy 3177 AAACAGCTGAATGCTGCTGAACCAAAAGCAAAAGTAAAGCAAAAGTAAAGTGGCG 3236
 Db 3226 AAACAGCTGAATGCTGCTGAACCAAAAGCAAAAGTAAAGCAAAAGTAAAGTGGCG 3285
 Qy 3237 TCAAAAAGAC-----AGTGTTCCTGATCCCTGCTGATCAAAAGCTGT---C 3284
 Db 3286 TCAAAAAGACATGAGAGAGCGTTTCTGATACCCGCTGATCAAAAGCTGTAAAC 3345
 Qy 3288 GCATTAAAGCGGCACTTGAATGATGCGCCAGCAAGCAATGCGAAAGATGCTCTA 3344
 Db 3346 GATTAAAGCGGCACTTGAATGATGCGCCAGCAAGCAATGCGAAAGATGCTCTA 3405
 Qy 3348 GATTAAGCGGCACTTGAATGATGCGCCAGCAAGCAATGCGAAAGATGCTCTA 3401
 Db 3402 GATTAAGCGGCACTTGAATGATGCGCCAGCAAGCAATGCGAAAGATGCTCTA 3461
 Qy 3466 AATAGTCCGTTATCGAGTGTCTGCAACGTAATAGTATGCTTCGTTCAAGATGA 3525
 Db 3467 TTAATGCTCTTTTGTAGATCAGCAATTCGCGGTGAGCAAAATTCAGCAAGAT 3521
 Qy 3526 TTGATGCTCTTTTGTAGATCAGCAATTCGCGGTGAGCAAAATTCAGCAAGAT 3585
 Db 3522 AAAAGCGTATGATCTGATGCGTTCCGCTTATTCAGCAAGCAAAAGCAATTAAGT 3581

Db 3586 AAAAGCGTATGATCTGATGCGTTCCGCTAT---CAGCAAGAAAGCAATTCGCT 3642
 Qy 3582 CAATTTGGGTGGAAAGACCTTGATTAATGACCAATTTGGGCGAGTTTTCGATAC 3641
 Db 3643 CAATTTGGGTGGAAAGACCTTGATTAATGACCAATTTGGGCGAGTTTTCGATAC 3702
 Qy 3642 CATTGATTAATCTCTTGAAGACAGTTAAATTCACGCGCATTAACGATGATGCG 3701
 Db 3703 CATTGATTAATCTCTTGAAGACAGTTAAATTCACGCGCATTAACGATGATGCG 3762
 Qy 3702 GGTTCGCAATATCAATGATGAGGCAATTAATTTGTTGTAAGTGGCAAGGCAATC 3761
 Db 3763 GGTTCGCAATATCAATGATGAGGCAATTAATTTGTTGTAAGTGGCAAGGCAATC 3822
 Qy 3762 AGTGGATTAATGCTGTAAGCAAAAGCCGAAATTTCAATGCAAAAGGCAATTAAT 3821
 Db 3823 AGTGGATTAATGCTGTAAGCAAAAGCCGAAATTTCAATGCAAAAGGCAATTAAT 3882
 Qy 3822 GCGTGAATGCAATGATCAGTCCGTTTGAAGCAATGCGCATTCAGCCCTTATTTGGA 3881
 Db 3883 GGTGGAATGCAATGATCAGTCCGTTTGAAGCAATGCGCATTCAGCCCTTATTTGGGT 3942
 Qy 3882 GTTAATGCTATTTATGTAAGGTAATTAATTCATTCAGCAAGTGAAGTGAAGC 3941
 Db 3943 GTTAATGCTATTTATGTAAGGTAATTAATTCATTCAGCAAGTGAAGTGAAGC 4002
 Qy 3942 CTTAGCCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4001
 Db 4003 CCGACCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4062
 Qy 4002 ACAGATTAATCAGCTTACGCTTATTTCTGATTAATGATGATGATGATGATGATG 4061
 Db 4063 ACAGATTAATCAGCTTACGCTTATTTCTGATTAATGATGATGATGATGATGATG 4122
 Qy 4062 AACGTAACCAACGTAATTTACGCTGTAACCAACGTAATTTACGCTGTAACCAAC 4121
 Db 4123 AACGTAACCAACGTAATTTACGCTGTAACCAACGTAATTTACGCTGTAACCAAC 4182
 Qy 4122 AAAGAGTGGATTTAAAGCAGAAATTTACATTCGAATTCGCTTATTCGTA 4181
 Db 4183 AAAGAGTGGATTTAAAGCAGAAATTTACATTCGAATTCGCTTATTCGTA 4242
 Qy 4182 TCTCAAGGTTGCAATCTGCGCAACGCAAAATGATGAGGCTGATGATGATG 4241
 Db 4243 TCTCAAGGTTGCAATCTGCGCAACGCAAAATGATGAGGCTGATGATGATG 4302
 Qy 4242 TAA 4244
 Db 4303 TAA 4305

RESULT 5
 US-10-329-960-1/c
 ; Sequence 1, Application US/10329960
 ; Publication No. US2003009277A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleischmann et al.
 ; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
 ; TITLE OF INVENTION: Thereof, and Uses Thereof
 ; FILE REFERENCE: P186P1
 ; CURRENT APPLICATION NUMBER: US/10/329,960
 ; CURRENT FILING DATE: 2003-01-02
 ; PRIOR APPLICATION NUMBER: US 09/643,990
 ; PRIOR FILING DATE: 2000-08-23
 ; PRIOR APPLICATION NUMBER: US 08/487,429
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: US 08/426,787
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1630121
 ; TYPE: DNA

ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (40868)..(40868)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102596)..(102596)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

Query Match	69.98;	Score 3019.8;	DB 15;	Length 1830121;
Best Local Similarity	83.08;	Pred. No. 0;		
Matches 3641; Conservative	2;	Mismatches 650;	Indels 93;	Gaps 14

Dh	278107	GATTATTCATTAAGCTCGTTTACATTAATTTGTACCTGAAGCTGAACCTGTGGGTATGACA	278048
Qy	5357	TCGATATGATGCGACGACTCTATTTCAGATAGAAACAAAATATCCAGAGCTGTTCATTC	556
Dh	278047	ACAAATATGATGAAAGAAAGTATATGCTGATATAGAGAAACTATCTGAGCGTTAGATATA	277988
Qy	557	GGCTTCGACGGCAGTTTGGGAAATGATCAAGCAAAAGCG-----	633
Dh	277987	GGCTCGAGACGCTGATATGGCGTACAGATTAAGATGAAGAAAGAAATGTACATATGTTCA	277928
Qy	640	-ACCAAGTGGCGGTGCATATCATATATGACAGCTGCATACACACATCAGCTGGA	638
Dh	277927	TATTAATGTCAGGTGATATGTTGTTATCTTACTCGACAGAAATCCCATCTCAGAGTGA	277868
Qy	699	GCAGGTATGATATTGTGATATTTGGAGCGCATGTTGTAAAGCGGAGAAATATSTTCA	758
Dh	277867	AATGTATATGTAACGTCAATCTTAGTGTAATATGTATACCTCATCATATATGTTCA	277808
Qy	759	TTACCATTCAGGCTCAAAAGGGGACAGTGGTCTTCGAGTTTATATATGATCTGAA	818
Dh	277807	TTACCAACGGGTGTTTAAAGGCGATAGCGGTTCGCAATGTTATTTATGATCGAAG	277748
Qy	813	AAACAAAATGGTTATATATGAGATATACGGGAAGCAACCTTTGTAGGCAAGAA	878
Dh	277747	AAGAAACAATGGCTTATTAATATGCTGATATACAACTGGGCATCTTTTTCGGAAGAGT	277688
Qy	879	AATGGGTTCATATGTTGGTCCAAATCTTATTTGAT-----GAATTTTC	933
Dh	277687	AATGGGTTCAGTTATATAGTGTAAGATAGTTTATATATGAAAGTTCTGGGTTGATACC	277628
Qy	924	GAAGAGATTTCATACATCATCTTACACCGGACGTGTATATGAGATGATACAAATTAGT	983
Dh	277627	CTTAGTGTTTTTCACGCTATATCTCCCAATTAAGACATATATCTTGTATCAAAAT	277568
Qy	984	GAATATGATATGTCAGGGGTCTATCTACTCAGAAATCAGAAATACCATCAGAAATTAA	1043
Dh	277567	AATGATGTACAGTTAAATTAATTTACTTACAGCTAGTAAAGATGGCTTTAAACAAA	277508
Qy	1044	ATTACGTAGCAAAATATGATTTACT-----TTGAAAGAGAGTAAAGTTGAT	1094
Dh	277507	TCGAAGTGAACATGTGAAGTATTTATTCATCTGTTAAATCAACCGCTAAAGACAT	277448
Qy	1095	AATCTTAGTATGACGACCTTAATTTATTTCTCCAGCTTTAAACAATGAGAAACGTA	1154
Dh	277447	GTTAAGCAGCGAGGCTAATATTTTACCAACAGAAATGAAATGTGAAAAAAATTT	277388
Qy	1155	TATTTATGATCAAAAGAGATCATATCTTCGACATGACATTAACCAAGGCGCG	1214
Dh	277387	TACTTCGGCACAGGAAGAAAGAACTTTACATGAAATTAATTAATCAAGTGTCT	277328
Qy	1215	GGTGTCTTATTTTGAAGGTTATTTACAGTATCTCCAA---ATTCTAACAACTGG	1271
Dh	277327	GGTGAATATCTTTGAGAGTATTTTCTGTGAAAAAGCAACAAATATATATATCTGG	277268
Qy	1272	CAGAGCTGSCATCATCTTATGTAATATGACACCGTTACTTGGAAATGAAATGCGCTG	1331
Dh	277267	CAGAGTGCAGGATATCTATGGAACAATGCAACGTGTGATGAAAGTCAATCTCT	277208
Qy	1333	GAACTGATGACTCTTCAAAATGTGTAAAGAACTTCGACGTTCAAGCCAAAGGGAA	1391
Dh	277207	GAATATATCTGTTTCTAAATATGTGATAGGCACTTTTATGCAATGTGATAGGAAAG	277148
Qy	1392	AATTAAGTTTCATCAGCGT-AGCGCATGTTAAATGATTTTGTGACGACGAGACGAT	1451
Dh	277147	AATTTAGGAAGTTTAATGGGGGTAAACGGCAAAATGCTTTCTAGATCAACAAGACAGTGA	277088
Qy	1452	CAGGCAACAAACAGCTTTAGTGAATTTGGCTTGTAGTGGCGACAGGAACTGTTCA	1511
Dh	277087	GGCGGTCAAAACAGGCTTTCAAGAAAGTGGCAATGTATAGGGTGCAGACAGGTTCA	277028
Qy	1512	TTTAACGATGATTAACAATTTGATACCAATAAATTTATTTTCGCTT-GGTGGTGGTGC	1571
Dh	277027	TTAATATGATCAGATCAAGTTATCTTACATATCTATTTTCGATTTTCGGTGGTGGTGC	276968

Db 274828 TATCAATGGGGGCAATTACATTTGGTGAATGAGGTGGGGAATAGTCCAGTAAA 274769
Qy 3774 ATGGCTGAAGACAAAGCCGAAATTCATCCGAAAGGAAATTAATGCGTGAATGCA 3833
Db 274768 ATGGCTGAAGACAAAGCCGAAATTCATCCGAAAGGAAATTAATGAGTGAATGCA 274709
Qy 3834 AGTATACATTCGCTTATAGGCAATTCGCGATTCAGCCTTATTTGGAGTAAATGCTAT 3893
Db 274708 AGTATACATTCGCTTATAGGCAATTCGCGATTCAGCCTTATTTGGAGTAAATGCTAT 274649
Qy 3894 TTTATGACGTAAGAAATTAATCAATCTGAGAGAGAGAGTGAAGAGCTTACGCTTGCA 3953
Db 274648 TTTATGACGTAAGAAATTAATCAATCTGAGAGAGAGAGTGAAGAGCTTACGCTTGCA 274589
Qy 3954 TTTATGACGTAAGAGTGGCATTCGAGTGAATTAATCAATTCCTCCGACATTAATATTC 4013
Db 274588 TTTATGACGTAAGAGTGGCATTCGAGTGAATTAATCAATTCCTCCGACATTAATATTC 274529
Qy 4014 AGCGTTAAGCTTATTTCTTCGTCATTAATGTTGATGTTTCAAGCTAAGCTAACAAACC 4073
Db 274528 AGCGTTAAGCTTATTTCTTCGTCATTAATGTTGATGTTTCAAGCTAAGCTAACAAACC 274469
Qy 4074 AGGTAATCTCAGCGTGTGCAACCAATTTGCAAGCTTAATGGCAAAAAGAGTGGGA 4133
Db 274468 ACTGTAATGACACATGTTGCAACCAATTTGCGGCGTATATGGCAAAAAGAGTGGGA 274409
Qy 4134 TTAAGGCAAGAAATTTACATTTCCAAATTTCCGCTTAAATCTCAAAATCTCAAGTCA 4193
Db 274408 TTAAGGCAAGAAATTTACATTTCCAAATTTCCGCTTAAATCTCAAAATCTCAAGTCA 274349
Qy 4194 CAATCGGCAAAACAGCAAAATGAGGCGTGAATTTGAGCTATGCTTGAATTAATCAACA 4253
Db 274348 CAATCGGCAAAACAGCAAAATGAGGCGTGAATTTGAGCTATGCTTGAATTAATCAACA 274289
Qy 4254 TAAATTTATGCTTATTAATTAACAGAGTGGTCAATCAATCCCACTTTTATTC 4313
Db 274288 TAAATTTATGCTTATTAATTAACAGAGTGGG-----GAGATCCCACTTTTATTC 274234
Qy 4314 AATAAT 4319
Db 274233 AATAAT 274228

RESULT 6
US-10-329-670-1/c
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: PH186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36536)..(36536)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40838)..(40838)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)

```

1 NAME/KEY: misc_feature
2 LOCATION: (145058)..(145058)
3 OTHER INFORMATION: n equals a, t, g or c
4 FEATURE:
5 NAME/KEY: misc_feature
6 LOCATION: (145171)..(145171)
7 OTHER INFORMATION: n equals a, t, g or c
8 FEATURE:
9 NAME/KEY: misc_feature
10 LOCATION: (145942)..(145942)
11 OTHER INFORMATION: n equals a, t, g or c
12 FEATURE:
13 NAME/KEY: misc_feature
14 LOCATION: (147197)..(147197)
15 OTHER INFORMATION: n equals a, t, g or c
16 FEATURE:
17 NAME/KEY: misc_feature
18 LOCATION: (150841)..(150841)
19 OTHER INFORMATION: n equals a, t, g or c
20 FEATURE:
21 NAME/KEY: misc_feature
22 LOCATION: (152500)..(152500)
23 OTHER INFORMATION: n equals a, t, g or c
24 FEATURE:
25 NAME/KEY: misc_feature
26 LOCATION: (152530)..(152530)

```

Query Match	69.9%;	Score 3019.8;	DB 16;	Length 1830121;	
Best Local Similarity	83.0%;	Pred. No. 0;			
Matches 364;	Conservative	2;	Mismatches 650;	Indels 93;	Gaps 14;

QY		1	CAATGTCGTACGA-GAATTTTAAAGAAAATTACTAATTAATAAACCT	59
Ds	278567		TCAATGTCGTACCGACGTAATTTTTTAATGAAAAATTACTAATTAATAAACAT	278528
QY		60	ATGAAAAAACTGATTAATTCGTCTTAATTTTAAACCGCTTGATTCATTGGGATAGA	119
Ds	278527		ATGAAAAAACTGATTAATTCGTCTTAATTTTAAACCGCTTGATTCATTGGGATAGA	278468
QY		120	TCCGACGCCTGGGCTGGTCAACATTAATTTTGGGATGATTACCATATATTCGTATTTT	179
Ds	278467		TCCGACGCCTGGGACGGTCACTTAATTTTGGGATGATTACCATATATTCGTATTTT	278408
QY		180	GCCGACAATAAAGGAAAGTTCACAGCTTGGGGCTCAAAATATTAAGGTTTAAACAAACA	239
Ds	278407		GCCGACAATAAAGGAAAGTTCACAGCTTGGGGCTCAAAATATTAAGGTTTAAACAAACA	278348
QY		240	GGGCAATTAAGTTGGCACATTAATGACAAAGCCCCGATGATTATTTCTGTAGTGTCA	299
Ds	278347		GGGCAATTAAGTTGGCACATTAATGACAAAGCCCCGATGATTATTTCTGTAGTGTG	278288
QY		300	CCTPAACGGGCTGGCAGCCTTGTTGTAATAATATATTTGTGAGCGTGGACATPACGA	359
Ds	278287		CCTPAACGGGCTGGGCGATTATGAGCGCATAGATATTTGTAGCGTGGACATPACGCC	278228
QY		360	GGATATPACGATGTTGTAATTTGGTGCAGAGGAAAACAACCAGATCAACATCGTTTACT	419
Ds	278227		GGATATPACGATGTTGTAATTTGGTGCAGAGGAAAGAACCCGATCAGCACGCCCTTACT	278168
QY		420	TATPACATGTTAAACGAAATAACTACAAA---AAAGATATATTACATCCTTATTAAGAAC	476
Ds	278167		TATPACATGTTAAAGAAATAATTAATTAACATCTGGGAGAGAAAGCAATCCTTATGATGA	278108
QY		477	GATTACCATATCCACGATTAGCTAAATTCCTTAGAGAGGGCTCCCAATATGATATGACT	536
Ds	278107		GATTATCATATGCTCGTTGTTACATAATTTTSTAATCGAAGCTGAACCTGGGGTATGACA	278048
QY		537	TCCGATATPAAATGACGACTTATTCACATATGAAACAAATATCCAGACGCTGTTGCTATC	596
Ds	278047		ACCAATATGAAAGAAAGATATATGCTGATAGSAGAACTA-CCGAGCGCTGTAGASTATA	277988
QY		597	GAGCTCTGAGCGGACGTTTGGGGAATATCATAGACAAAGSCG-----	639

QY 2766 AATTAATTAATCAATGACGCTGAGGCGATTACATATTATCTGTTCGACACAGAGC 2825
 DB 275767 AATTAATTAATTAATCAATGACGCTGAGGCGATTACATATTATCTGTTCGACACAGAGC 275708
 QY 2826 AAGAAACCCGAAACCTTGAAGCAATTAATCTTGTGAAAGCAAGATTAATCAACGTTA 2885
 DB 275707 AAGAAACCTGAGACCTTGAAGCAATTAATCTTGTGAAAGCAAGATTAATCAACGTTA 275448
 QY 2886 TCGATTAAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2945
 DB 275647 TCGATTAAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275588
 QY 2946 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3005
 DB 275587 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275528
 QY 3006 CACAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3065
 DB 275527 CACAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275468
 QY 3066 CCGATCTGCTAAACCAACAGGCTGACCAAAAGTGGCTCAAGAGAGCGAGCA 3125
 DB 275467 CCGATCTGCTAAACCAACAGGCTGACCAAAAGTGGCTCAAGAGAGCGAGCA 275417
 QY 3126 GCGTTCTGATACCCCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAAGCT 3185
 DB 275416 GCGTTCTGATACCCCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAAGCT 275357
 QY 3186 GAACTGATCTGAAACCAACAAAGATTAAGGCAAAACAAAGAGTGGCTCAAAAGCA 3245
 DB 275356 GAACTGATCTGAAACCAACAAAGATTAAGGCAAAACAAAGAGTGGCTCAAAAGCA 275297
 QY 3246 GCGTTCTGATACCCCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAAGCT 3293
 DB 275296 GCGTTCTGATACCCCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAAGCT 275237
 QY 3294 GCGGCACTTGAAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3353
 DB 275236 GCGGCACTTGAAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275177
 QY 3354 GAAAGCGAAACCAACCAACAAAGATTAAGGCAAAACAAAGAGTGGCTCAAAAGCA 3413
 DB 275176 GAAAGCGAAACCAACCAACAAAGATTAAGGCAAAACAAAGAGTGGCTCAAAAGCA 275126
 QY 3414 TCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3473
 DB 275125 TCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275066
 QY 3474 TTTGTGATCAAGCAACATCTGCGTGTGACAAATTAATTAATTAATTAATTAATTAAT 3533
 DB 275065 TTTGTGATCAAGCAACATCTGCGTGTGACAAATTAATTAATTAATTAATTAATTAAT 275006
 QY 3534 GATTCTGATGCGTTCGCTGCTTATCAAGCAGAGAAACCAATTAATTAATTAATTAAT 3593
 DB 275005 GATTCTGATGCGTTCGCTGCTTATCAAGCAGAGAAACCAATTAATTAATTAATTAAT 274949
 QY 3594 CAAAAAGCTTGAAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3653
 DB 274948 CAAAAAGCTTGAAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274889
 QY 3654 ACCTTGAATGAACAGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3713
 DB 274888 ACCTTGAATGAACAGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274829
 QY 3714 TATCAATGAGGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3773
 DB 274828 TATCAATGAGGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274769
 QY 3774 ATGCGTGAAGAAACCAACCAACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3833
 DB 274768 ATGCGTGAAGAAACCAACCAACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 274709

QY 3834 AGTATCAGTTCCTGTTAGGCAATGGGCAATCACCCTTAATTTAGAGTATAGCTAT 3893
 DB 274708 AGTATCAGTTCCTGTTAGGCAATGGGCAATCACCCTTAATTTAGAGTATAGCTAT 274649
 QY 3894 TTTATTAAGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3953
 DB 274648 TTTATTAAGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274589
 QY 3954 TTTATTAAGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4013
 DB 274588 TTTATTAAGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274529
 QY 4014 ACCTTGAAGCTTATTTCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 4073
 DB 274528 ACCTTGAAGCTTATTTCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 274469
 QY 4074 ACGTAAATCTACAGGCTGTGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 4133
 DB 274468 ACGTAAATCTACAGGCTGTGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 274409
 QY 4134 TTTAAAGCAAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4193
 DB 274408 TTTAAAGCAAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274349
 QY 4194 CAACTGGCAAAACGCAAAATGAGGCGTGAATTTGGGCTATCGTGTAAATTAATTAATTAAT 4253
 DB 274348 CAACTGGCAAAACGCAAAATGAGGCGTGAATTTGGGCTATCGTGTAAATTAATTAATTAATTAAT 274289
 QY 4254 TAAATTAATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4313
 DB 274288 TAAATTAATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274234
 QY 4314 AATTAAT 4319
 DB 274233 AATTAAT 274228

RESULT 7
 US-10-080-505-12
 ? Sequence 12, Application US/10080505
 ? Publication No. US2003073166A1
 ? GENERAL INFORMATION:
 ? APPLICANT: St. Gene, Joseph W.
 ? TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ? FILE REFERENCE: A-59941-1/RT/DCV/DHR
 ? CURRENT FILING DATE: 2002-02-22
 ? PRIOR APPLICATION NUMBER: US 08/296,791
 ? PRIOR FILING DATE: 1994-10-25
 ? PRIOR APPLICATION NUMBER: US 09/839,996
 ? PRIOR FILING DATE: 2001-04-20
 ? NUMBER OF SEQ ID NOS: 58
 ? SOFTWARE: PatentIn version 3.1
 ? SEQ ID NO 12
 ? LENGTH: 5245
 ? TYPE: DNA
 ? ORGANISM: Haemophilus influenzae
 ? FEATURE:
 ? NAME/KEY: CDS
 ? LOCATION: (430)..(4740)
 ? OTHER INFORMATION:
 US-10-080-505-12

Query Match 69.2%; Score 2990.4; DB 15; Length 5245;
 Best Local Similarity 82.0%; Pred. No. 0;
 Matches 3660; Conservative 0; Mismatches 636; Indels 168; Gaps 12;

QY 1 TCAATAGCTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 59
 DB 370 TCAATAGCTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 429
 QY 60 ATGAAAAAAGCTATTTGCTTAAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 119

Db 430 ATGAAAAAACTGATTTTCGTGTAACCTTTTAAACCGCTTGCACTTTCATTTAGGAGTACTA 489
 Qy 120 TCGCAAGCTGGGCTGGTCACTTATTTTGGATGATTAACAATATATGATGATTT 179
 Db 490 TCGCAAGCTGGGCTGGTCACTTATTTTGGATGATTAACAATATATGATGATTT 549
 Qy 180 GCCGAAATTAAGGGAAGTTCAACGTTGAGGCTCAAAATATTAAGTTTATACAAACA 239
 Db 550 GCTAGAAATTAAGGGAAGTTTCACTTTGGGCTTAAAAATATTAAGTTTATACAAAG 609
 Qy 240 GGGCAATTAAGTTGGCAATCAATGCAAAAGCCCGATGATTAATTTTCTGAGTCA 299
 Db 610 GGGCACTTAAGTTGGCAATCAATGCAAAAGCCCGATGATTAATTTTCTGAGTCA 669
 Qy 300 GGTAAAGGCTGGGCTGGTGGAAATCATATATTTGAGCGGTGCACTAAGCTA 359
 Db 670 GGAATGAGGCTGGGCTGGTGGAAATCATATATTTGAGCGGTGCACTAAGCTA 729
 Qy 360 GGAATTAAGGCTGGTGGTGGTGGGGAAGCAACCCGATCAACATGCTTACT 419
 Db 730 GGAATTAAGGCTGGTGGTGGGGAAGCAACCCGATCAACATGCTTACT 789
 Qy 420 TATAGATTTGAAAGGAATATCAAAAAGATATTAATTAATTAATTAATTAATTAAT 479
 Db 790 TATCAAAATGAAAGGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 849
 Qy 480 TACATATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 539
 Db 850 TATCAATGCTGCTGGTGGCAAAATTTGATGAGCTGCAACCAAGAAATGATGAGCT 909
 Qy 540 AATATGAATGCAATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 599
 Db 910 AATATGAATGCAATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 969
 Qy 600 TCTGAGGAGGCTTTGGGGAATGATCAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAG 644
 Db 970 AAGAGTAAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1029
 Qy 645 GTTCCGCTGAT 704
 Db 1030 CTTCGATGCTTAT 1089
 Qy 705 AATGATATGCTATTTGGGAGGAGGATGTTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGG 764
 Db 1090 AAGGATATGCTATTTGGGAGGAGGATGTTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1149
 Qy 765 AATGAGGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 824
 Db 1150 AGGAGTGTTCGTTGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1209
 Qy 825 AATGATTAATTAATGAGATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 884
 Db 1210 AATGATTAATTAATGAGATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1269
 Qy 885 TTTCAATTTGCTGCAATCTTATTTGATGAAA---TTTTCGAAAGATTTTACATCA 941
 Db 1270 TTTCAATTTGCTGCAATCTTATTTGATGAAA---TTTTCGAAAGATTTTACATCA 1329
 Qy 942 TCACTTAAACCCGAGCTGTAATGAGGTAACAATTAAGGAAATTAAGTATAGTCAAG 1001
 Db 1330 AATTTTAAAGCCAGAGATTAAGGTAACAATTAAGGAAATTAAGTATAGTCAAG 1052
 Qy 1002 GGGCTTAAATC-----AGAAATCAGAAATACCAATTAAGTATAGTCAAG 1052
 Db 1390 GGTACGGTACTCAAGATGAAGAAAGTATAGTCAATTAAGTATAGTCAAG 1449
 Qy 1053 GGAATATGAGTTTACCTTGAAGAGAGATTAAGTCAATTAAGTATAGTCAAG 1112
 Db 1450 CAGTATATTAAGGATTAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 1509
 Qy 1113 CCAATATTTATCTCAAGCTTAAACATGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 1172
 Db 1510 GAAATGCTTAAACCAAGATTAAGTATGAAATATTTTACCTTTGGGAGTCAAG 1569

Qy 1173 CAAGATCATATATCTTGCATCTGACATTAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1232
 Db 1570 AAGGATCTTAAATTAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1629
 Qy 1233 GGTATTTTAAAGATATCTCAAAATTTCAACCAATTTGCAAGGAGGAGGAGGAGGAGGAGGAG 1292
 Db 1630 GGTATCTTAAAGGATATCTTCAAGAAATTAATCAAACTTTGCAAGGAGGAGGAGGAGGAGGAG 1689
 Qy 1293 AGTAAATATGACCGTTACTTGAAGATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1352
 Db 1692 GGTAAAGACAGATCTTACTTGAAGATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1749
 Qy 1353 ATTGTAAGGACATTTGCAAGCTTCAAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1412
 Db 1750 ATTGTAAGGACATTTGCAAGCTTCAAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1809
 Qy 1413 GGGATGTAAGGATCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1472
 Db 1810 GGTAGCGGCAAGGATCTTGAATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1869
 Qy 1473 AGTAAATTTGGCTGTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1532
 Db 1870 AAGAAAGTTGGATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1929
 Qy 1533 GATACGATTAATTTTATTTTGGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1592
 Db 1930 GATCTTAACATTTTATTTTGGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1989
 Qy 1593 TTAACCTTTAAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1652
 Db 1990 TTAACCTTTAAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2049
 Qy 1653 ACTCAAGCCGCTTATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1709
 Db 2050 ACTCAAGCCGCTTATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2209
 Qy 1710 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1769
 Db 2110 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2169
 Qy 1770 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1829
 Db 2170 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2229
 Qy 1830 GGTACTTGTCTACTTCAAGTGTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1889
 Db 2230 GGTACTTGTCTACTTCAAGTGTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2289
 Qy 1890 AATCTATTTTCAAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1949
 Db 2290 ACTTATGTTTATGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2009
 Qy 1950 TCGAAATGAAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2059
 Db 2344 CCAAGAGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2109
 Qy 2310 AATTTTAAAGCTGAAATCTTCAAAATTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2169
 Db 2404 AATTTTAAAGCTGAAATCTTCAAAATTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2229
 Qy 2070 TCTTCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2289
 Db 2464 TCTTCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2349
 Qy 2130 CCAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2409
 Db 2524 CCAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2469
 Qy 2190 AAGTGAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2549
 Db 2584 ACTGTGATTTTAAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2609

Db 4789 --CCACCCTTTTATTCATATAT 481C

RESULT 8
US-10-080-505-16
Sequence 16, Application US/10080505
Publication No. US20030073166A1
GENERAL INFORMATION:
APPLICANT: St. Gene, Joseph W.
TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
REFERENCE: A-59941-1/RPT/DCF/DHR
CURRENT FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: US 08/296,791
PRIOR FILING DATE: 1994-10-25
PRIOR APPLICATION NUMBER: US 09/839,996
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 4828
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: CDS
LOCATION: (313)..(4548)
OTHER INFORMATION:
US-10-080-505-16

Query Match 68.8%; Score 2970; DB 15; Length 4828;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 358; Conservative 0; Mismatches 645; Indels 124; Gaps 10;

QY 1 TCATATGCTGTTACTA-GTATTTTATACGAAATTAATTAATAATCACT 59
Db 253 TCATATGCTGTTACTA-CGATTTTATACGAAATTAATTAATAATCACT 312

QY 60 ATGAAAAAAGTATTTTCTTTAATTTTAAACCGCTTCATTAAGGATAGTA 119
Db 313 ATGAAAAAAGTATTTTCTTTAATTTTAAACCGCTTCATTAAGGATAGTA 372

QY 120 TCGAAGCGTGGCGGTGCTCACTTATTTGGGATGATTAACAAATTAATGATTT 179
Db 373 TCGAAGCGTGGCGGTGCTCACTTATTTGGGATGATTAACAAATTAATGATTT 432

QY 180 GCGAGATTAAGGGAAGTTCACAGTTGGGCTCAAAATTAAGGTTATTAACAACA 239
Db 433 GCGAGATTAAGGGAAGTTCACAGTTGGGCTCAAAATTAAGGTTATTAACAACA 492

QY 240 GGGCATTTGTTGGCATCATGATGACAAAAGCCCGATGATTTTCTGATGTC 299
Db 493 GGAATTTGTTGGCATCATGATGACAAAAGCCCGATGATTTTCTGATGTC 552

QY 300 GGTAGCGGTGGCGGCTTGTGTAATCAATATTTGAGGCTGGCATTAAGTA 359
Db 553 GGAATTTGTTGGCATCATGATGACAAAAGCCCGATGATTTTCTGATGTC 612

QY 360 GATATATACAGATTTGATTTGGTGGAGGGAACAACCCGATCAACATGTTACT 419
Db 613 GGTATATACAGATTTGATTTGGTGGAGGGAACAACCCGATCAACATGTTACT 672

QY 420 TATATATTTTAAACGAAATTAACTAACAAAAGATTAATTAATCTTATGAGACAT 479
Db 673 TATATATTTTAAACGAAATTAACTAACAAAAGATTAATTAATCTTATGAGACAT 732

QY 480 TACCATATTCAGATTTACATAATTCGTTGAGAGGCTGCAATGATGATGATG 539
Db 733 TACCATATTCAGATTTACATAATTCGTTGAGAGGCTGCAATGATGATGATG 792

QY 540 AATATGATGGCAGTACTTATTCAGATAGAACAAATATTCAGAAAGTGTTCATGCG 599
Db 793 GATATGATGGCAGTACTTATTCAGATAGAACAAATATTCAGAAAGTGTTCATGCG 852

QY 600 TCTGACCGCAGTTTGGGAAATGATCAACAAAGCGAACCAAGTTGCGGTGATAT 659
Db 853 TCGGGTGGCAGTTTGGGAAATGATCAACAAAGCGAACCAAGTTGCGGTGATAT 912

QY 660 CATATCTGACAGTGGCAATACACATACAGTGGAGCAGTAAATGATGAT 719
Db 913 CATATCTGACAGTGGCAATACACATACAGTGGAGCAGTAAATGATGAT 972

QY 720 TTGGAGGCGGTGTTAAACCGGGAATATGTCCTTAACGATTTGAGGCTCAAG 779
Db 973 CTGACGCGGATGTCGACCAAGCGGCAATTAAGCGCCATTCATTTGAGGCTCAAG 1032

QY 780 GGGGACAGTGTCTCCGATGTTATTTATGATGCTGAAAAACAAATAGTAAATAT 839
Db 1033 GGGGACAGTGTCTCCGATGTTATTTATGATGCTGAAAAACAAATAGTAAATAT 1092

QY 840 GGAATATTAAGGAAAGGCAACCTTTGAAAGCAAAAGAAATGATTTCAATGTTGCG 899
Db 1093 GGGATATTAAGGAAAGGCAACCTTTGAAAGCAAAAGAAATGATTTCAATGTTGCG 1152

QY 900 AATCTTATTTTGTAAATTTTGAAGAGATTAATCAATGATCACTTACACCGAGCT 959
Db 1153 AATCTTATTTTGTAAATTTTGAAGAGATTAATCAATGATCACTTACACCGAGCT 1222

QY 960 GGTAAATGAGGTACACATATTAAGTAAATGATGATGTCAGGAGTATTAATCAAGAA 1019
Db 1213 GGTAAATGAGGTACACATATTAAGTAAATGATGATGTCAGGAGTATTAATCAAGAA 1272

QY 1020 TCGAATATTCATCAAAATTAATTAATGATGCAATATGATGATGATGATGATG 1079
Db 1273 TCGAATATTCATCAAAATTAATTAATGATGCAATATGATGATGATGATGATG 1332

QY 1080 AAGATTAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1139
Db 1333 AAGATTAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1386

QY 1140 AATGAGAAAGCTATATTTATTAAGATCAAAACCAAGATCAATTAATCTGATGAT 1199
Db 1387 AATGAGAAAGCTATATTTATTAAGATCAAAACCAAGATCAATTAATCTGATGAT 1446

QY 1200 ATTAACTGAGGCGGCTGCTTATTTTGAAGGATTTTAAAGATTTTCAATTTCT 1259
Db 1447 ATTAACTGAGGCGGCTGCTTATTTTGAAGGATTTTAAAGATTTTCAATTTCT 1506

QY 1260 AACCAACTGAGCAAGAGCTGCAATGATGATGATGATGATGATGATGATGATG 1319
Db 1507 AACCAACTGAGCAAGAGCTGCAATGATGATGATGATGATGATGATGATGATG 1566

QY 1320 GTAAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1379
Db 1567 GTAAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1626

QY 1380 GCGAAAGGGAATTAAGATTTGATGATGATGATGATGATGATGATGATGATG 1439
Db 1627 GCGAAAGGGAATTAAGATTTGATGATGATGATGATGATGATGATGATGATG 1686

QY 1440 CAGGCAAGCATCAAGGCAACCAAGCTTTAGTAAATTTGCTGTTGAGCGGCA 1499
Db 1687 CAGGCAAGCATCAAGGCAACCAAGCTTTAGTAAATTTGCTGTTGAGCGGCA 1746

QY 1500 GGGACTGTTAATTAAGCATTAACATTTGATGATGATGATGATGATGATGATG 1559
Db 1747 GGGACTGTTAATTAAGCATTAACATTTGATGATGATGATGATGATGATGATG 1806

QY 1560 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1619
Db 1807 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1866

QY 1620 GACGAGGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1679
Db 1867 GACGAGGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1926

QY 1680 GGGAAAGAAAGCATTTGTTACTTAATGAAATATTAATTAATTAATTAATTA 1739

Db 1522 ACTTTAAGATTGAAATTAATTAATCAAGTGCAGTGGATTTGATTTTGAAGTAAT 1587-
Qy 1239 TTATACAGATATCCAAAT---TCTAACCAACTTGGCAAGAGCTGGATACATGTAAGT 1295
Db 1582 TTGTGTTAAAGGCATCAAAATATATTAATCTTGGCAAGTGGAGGTTCTGTGTA 1641
Qy 1296 GAAATAGCACCTTACCTTGGAAAGTAATGGCTGGAAATGATGCACTTTCTAAAT 1355
Db 1642 GAAGAAATCTGTATTAATGGAGGTGATATTCAGAAAGGCGCTATTCAAAAT 1701
Qy 1356 GGTAAAGGACATTCACCTTCAAGCCAAAGGGGAAAATAAGGTGGATCAAGAGGC 1415
Db 1702 GGGCTGGACATCTTACTTGTATGTATTAAGGAAAACCTTGAAGCTGAGTGTGCT 1761
Qy 1416 GATGTAAAGTACTTTTGGACAGAGGCAAGATCAAGCAACAAACAGCTTAAAT 1475
Db 1762 AAGGTTGTGTGTAGTATCAACAGCAGATGATTCAGGTCAAAAACAGCTTAA 1821
Qy 1476 GAAATGTGCTGTAGCGGCAAGAGGACTGTTCAATTAACGATGATTAACATTTAT 1535
Db 1822 GAAGTGGCATGTATAGTGTAGAGCTACGTTCAATTAATAGTGCATCAAGTGTAT 1881
Qy 1536 ACCGATTAATTTATTTGGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1595
Db 1882 CCTAACATATTTATTTGGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1942
Qy 1596 ACCTTAAACGATCCAAATATCGAGAGAGGGGCAATGATTTGAACATTAACAAT 1655
Db 1942 ACCTTGAAGGTATCCAAATATCGAGATGAAGGGGCAATGATTTGAACCAACGCTCT 2001
Qy 1656 CAAGCGCTATGTCTATTTCTGTGGAAAGAAAGCATTTCTACTTAATGAATATAT 1715
Db 2002 CAAGCGCAATATATCAATTAAGGCAAGCAACATTAATTAATCA-----GAT 2049
Qy 1716 ATTAATTAACCTGATTAACGAAAGAAATTTGCCCAACGGTGTGTGTGTGTGTGTGT 1775
Db 2052 AGCAACCACTTACTTAATTAAGAAATTTATTCATTTAAAGCTGTGTGTGTGTGTGT 2109
Qy 1776 AAAAATTAACCAATGGGCGATTAACCTTATTAATTAACCAACCAAGAGTGTACT 1835
Db 2112 AAAGCTAAACCAATGGTGTGTATTAATGATTAATCAACGATTAATGAAGAAATAT 2169
Qy 1836 TTGCTACTTCAAGTGTGTATTAATTAAGGCGATTTACCAACAAAGGTAACATA 1895
Db 2170 TTGTGCTTTCTGGGGGCAAAATTTAAAGGCAATATCAAGCAAAATGGTGTGTGT 2229
Qy 1896 TTTTCAAGCGTAGACCAACGAGCGAGCTCAATCAATTAATTAAGGTTGTGTGTGT 1955
Db 2230 GTTTTATGT 2289
Qy 1956 ATGAAAGGTATACCAAGAGGGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2015
Db 2290 ATGAAAGGTATCCCAAGAGGGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2349
Qy 2016 AAAGCTGAAACTTCCAAATTAAGGCGGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 2075
Db 2350 AAAGCTGAAACTTCCAAATTAAGGCGGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 2409
Qy 2076 ATGAGGGAATTTGACAGTCAAGATTAATGCAAAATGCAATTTGTGTGTGTGTGTGT 2135
Db 2410 ATGAGGGAATTTGACAGTCAAGATTAATGCAAAATGCAATTTGTGTGTGTGTGTGT 2469
Qy 2136 CAACAAATACATTTGACAGGCTTCAAGATTTGAAGATTAAGATTTGTGTGTGTGT 2195
Db 2470 CAGCAAAATACATTTGACAGGCTTCAAGATTTGAAGATTTGAAGATTTGTGTGTGT 2529
Qy 2196 GATTAACGATTAAGGATTAATTTATTTATTAATTAAGCAAAATCAATCAATGTGTGT 2255
Db 2530 GATTAACGATTAAGGATTAATTTATTTATTAATTAAGCAAAATCAATCAATGTGTGT 2589
Qy 2256 AATTAACTGATTAAGCAAGGCAATGTTAAAGGTTTAAAGCAAAATCAATCAATGTGT 2315
Db 2590 AATTAACTGATTAAGCAAGGCAATGTTAAAGGTTTAAAGCAAAATCAATCAATGTGT 2649

Qy 2316 ACTTAACAAATCAAGCCATTTA---CTTAAGCAACATGCAACCCCAATGAGCAATAT 2375
Db 2650 ACTTAATATATCAAGCCATTTA---CTTAAGCAACATGCAACCCCAATGAGCAATAT 2709
Qy 2316 CGACTTCCGCAATTCACCTGCAAGGTGTATTAAGCAAACTGAAACGGTATGTGAT 2435
Db 2710 AAACCTTCAATCAAGCAATGCAAGGTGTGCAATGCAAACTGAAACGGTATGTGAT 2769
Qy 2436 TTAAGGATTTAGCTCAATTTTCTTTAAAAAACAGCCTTTTGTGCAACCAATTTAGGA 2495
Db 2770 TTAATGATTTGTGCAATTTTCTTTAAAAAACAGCCTTTTGTGCAACCAATTTAGGA 2829
Qy 2496 GACAAAGCAACAGTGTGAGGAAATTTGCAAACTTGTGCAAACTGAGCATATCA 2555
Db 2830 GGGAAAGCAACAGTGTGAGGAAATTTGCAAACTTGTGCAAACTGAGCATATCA 2889
Qy 2556 TTGCAAGATTTAAGCTTAATTAAGATGATCAAGTAAATCAAGCTTATCAAGTATG 2615
Db 2890 TTGCAAGATTTAAGCTTAATTAAGATGATGATGATGATGATGATGATGATGATGAT 2949
Qy 2616 TCAAAATTAAGGCA-----CGTGGCTGTATTAAGGCAAGCAACAGCATG 2669
Db 2950 TCAAAATTAAGGCAAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3009
Qy 2670 GCAAGACATGCTTTCAACATTTGACAGTAAATGATTAATTTAATTAATTAATTAATTA 2725
Db 3010 GCAAGACATGCTTTCAACATTTGACAGTAAATGATTAATTTAATTAATTAATTAATTA 3069
Qy 2730 CAATTTACTCATCTTTATTTGTGCTTAAAGCAATTAATTAATTAATTAATTAATTAAT 2789
Db 3070 CAATTTACTCATCTTTATTTGTGCTTAAAGCAATTAATTAATTAATTAATTAATTAAT 3129
Qy 2730 GAGGGGATTAATTAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2849
Db 3130 GAGGGGATTAATTAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3189
Qy 2850 TTAATTTGTGTGAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2909
Db 3190 TTAATTTGTGTGAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3249
Qy 2910 GAAATAGCACCTTATGAGT 2969
Db 3250 GAAATAGCACCTTATGAGT 3309
Qy 2970 TTCCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3029
Db 3310 TTCCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3369
Qy 3030 CAAGCAAGCAATTAAGAGCCAAACAGTGAACCGCTGTAACCAACCAACAGT 3089
Db 3370 CAAGCAAGCAATTAAGAGCCAAACAGTGAACCGCTGTAACCAACCAACAGT 3429
Qy 3090 GAGCCAAAGT 3149
Db 3430 GAGCCAAAGT 3480
Qy 3150 CAAGGCTGTGTGAAGGATTAAGGCAACCAAGTGAATGAGTGTGTGTGTGTGTGTGTGT 3209
Db 3481 CAAGGCTGTGTGAAGGATTAAGGCAACCAAGTGAATGAGTGTGTGTGTGTGTGTGTGT 3537
Qy 3210 AGTAAGGCAAAACCAAAAGGCGCTCAAAAGAGCAGTGTGTGTGTGTGTGTGTGTGTGT 3269
Db 3538 AGTAAGGCAAAACCAAAAGGCGCTCAAAAGAGCAGTGTGTGTGTGTGTGTGTGTGTGT 3597
Qy 3270 GATCAAGCTGT 3329
Db 3598 GATCAAGCTGT 3648
Qy 3330 GAAAAAGATGCTGT 3386
Db 3649 AAAAAAGATGCTGT 3708

RESULT 10
US-10-080-505-10
; Sequence 10, Application US/10080505
; Publication No. US20030073166A1
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/ST/DCF/DIR

Query Match	62.7%	Score 2709.6	DB 15	Length 4605
Best Local Similarity	79.4%	Pred. No. 0		
Matches 3401; Conservative	0	Mismatches 809	Indels 76	Gaps 13

[illegible]

QY 777 AAGGGGACAGTGTCTTCGAGTGTATTTATGATGCTGAAAAAAGAAATGTATTT 836
 Db 1136 TTGGGGATAGTGTCTTCGAATGTATTTATGATTAAGAAATGTAAGAAATGTTATTA 1195
 QY 837 AATGGATATTAAGGAAAGCAACCTTTTGAAGCAAGAAATGGCTTCAATGGTT 896
 Db 1136 AATGGCTGTATGAAAGAAATCTTATGCTGATGAAACAGCTATCAATTAACA 1255
 QY 897 CGCAATCTTATTTGATGATAATTTTGAAAGATTTACATCATCCTTTACACCGGA 956
 Db 1256 GAAAAATTTATTTTCAAGATTTCTTATTAACAGCATTTACAGCTAATTTTGGATCT 1315
 QY 957 GCTGTAAATGAGTGTACCAATTTAGTGAATGATTAATGCTGAGGCTC----- 1006
 Db 1316 AATGCTAATATAGATTTAATATAGGAGTGAACCAATGGAAGATGGCAACATGAA 1375
 QY 1007 --TATACCTCAAAATCGGAATTCATCGAAATTTAAATTTACCTTACCAATTTAGT 1064
 Db 1376 AGTACATTAATCTAAAAAGCTATTCAGCTGAACGAAATGAGGCTCTTATGATTAATAGC 1435
 QY 1065 TTACCTTTGAAAGAAAGATTA-----GTCTAATATCTGATATAGCGAATAT 1118
 Db 1436 CAATCTCATGATCTGAGATTAAGAAATGCGATGATATCTCTCTTATTAAGGCTCTAT 1495
 QY 1119 ATTATTTCTCAGCTTTAAACATGAGAAAGCTATATTTATGATCAAAAACAGAA 1178
 Db 1496 CCAATGCTCCAGCATTAATCATGAGAAAGATTTATCTTGGCATCAAGAAACAGAA 1555
 QY 1179 TCATTTATCTTGCACTGATTAACCAAGGAGGCTGTCTTATTTGAGGAT 1238
 Db 1556 ACTTAAATATGAAATTAATATTAATTAATCAAGGATGAGGATTTGATTTGAAGTAA 1615
 QY 1239 TTTACAGTATCTCAAT---TCTAACAACTTGCAGAGACTGCAATCATATGAT 1295
 Db 1616 TTGTGTGTAAGGAAATCAAAATTAATTAATCTTGCAAGGCTGTCTTCTGTGGA 1675
 QY 1296 GAAATGACACGTTACTTGAAAGTAAATGCGATGATGATGATCTTCTAAAT 1355
 Db 1676 GAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1735
 QY 1356 GGTAAAGAACTTGCAGCTTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 1415
 Db 1736 GGGCTGGAACCTTACTTGTAAATGTAAGGAAAGGAAAGGAAAGGCAAGGCTGAGT 1795
 QY 1416 GATGTAAAGTCAATTTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1475
 Db 1796 AAGGTTGTGTGTGATTAAGTAAAGCAGATGATGATGATGATGATGATGATGAT 1855
 QY 1476 GAAATGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1535
 Db 1856 GAAATGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1915
 QY 1536 ACCGATTAATTTTATTCGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1595
 Db 1916 CTTAAATATTTATTTGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1975
 QY 1596 ACCTTTAAAGTATCCAAATACGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1655
 Db 1976 ACCTTTGAAGGATCAAAATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2035
 QY 1656 CAAGCCGTAATGCTATTTACTGGAACGAAAGCAATGTTCTACCTTAATGAAATAT 1715
 Db 2036 CAAGCCGTAATTTAGATTTACGAGCAACGAACTATTAATTTCA-----GAT 2083
 QY 1716 ATTAATTAATCTGATTAAGAAAGAAATGCTAAGAGGTTGTGTGTGTGTGTGTGTGT 1775
 Db 2084 AGCAATCAATCTTAAATTAAGAAATGATGATTTAAAGGCTGTGTGTGTGTGTGTGT 2143
 QY 1776 AAAAATTAACCAATGAGGAGTAAACCTTATTTAATTAACCAACCAAGAGATGCT 1835
 Db 2144 AAGCTAAAGCAATGAGGCTTTAATGATGATTAATCAACGATTAATGAGCAAGAAATCAT 2203

QY 1836 TTGCTACTTTCAGGTGTACAAATTTAAAGGAGATTTACCCAAACAAAGATTAATTA 1895
 Db 2204 TTGTCTCTTTCGAGGGGCAAAATTTAAAGGCAATATCACCCAAAATGAGGATGTTA 2263
 QY 1896 TTTTTCAGGCTAGACGACACCGCAACGCTTACATCATTTTAAATTAAGGTTGCTAGAA 1955
 Db 2264 GTTTTATGAGTGTCTTCAACGCTCATCTTACATCATTTAAGAGAGAGATCTGCTAAC 2323
 QY 1956 ATGAAAGGATTAACCAAGGCGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2015
 Db 2324 ATGAAAGGATTAACCAAGGCGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2383
 QY 2016 AAGCTGAAATCTTCAATTTAAAGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2075
 Db 2384 AAGCTGAAATCTTCAATTTAAAGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2443
 QY 2076 ATTGAGGAAATTTGACAGTGTGCAATTAATGCAATGCAATTTGTGTGTGTGTGTGT 2135
 Db 2444 ATTGAGGAAATTTGACAGTGTGCAATTAATGCAATGCAATTTGTGTGTGTGTGTGT 2503
 QY 2136 CAACAAATACATTTGCAACGCTTCAATTTAAAGGAGAGTGTGTGTGTGTGTGTGTGT 2195
 Db 2504 CAGCAAAATACATTTGCAACGCTTCAATTTAAAGGAGAGTGTGTGTGTGTGTGTGT 2563
 QY 2196 GATTTACCGATTAACAAAGTATTAATTTAATTTAATTTAATTTAATTTAATTTAAT 2255
 Db 2564 GATTTACCGATTAACAAAGTATTAATTTAATTTAATTTAATTTAATTTAATTTAAT 2623
 QY 2256 AATTTACTGATTAATGCAACGCGAATTTAAAGGTTTAAAGGTTTAAAGGTTTAAAG 2315
 Db 2624 AATTTACTGATTAATGCAACGCGAATTTAATTTAATTTAATTTAATTTAATTTAAT 2683
 QY 2316 ACTTTACAAATTAACGCGCAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 2375
 Db 2684 ACTTTACAAATTAACGCGCAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 2743
 QY 2376 CCAATTTCCGCAATTAACGCGAATTTAATTTAATTTAATTTAATTTAATTTAATTTA 2435
 Db 2744 AATTTACTGATTAATGCAACGCGAATTTAATTTAATTTAATTTAATTTAATTTAAT 2803
 QY 2436 TTAACGATTTAGTCAATTTTCTTAAATTAATTTAATTTAATTTAATTTAATTTAAT 2495
 Db 2804 TTAACGATTTAGTCAATTTTCTTAAATTAATTTAATTTAATTTAATTTAATTTAAT 2863
 QY 2496 GCAAAAGGCAACAGGAGCTTGAAGAAATGCAAGTGAAGTGAAGTGAAGTGAAGTGA 2555
 Db 2864 GGGAAAGCAACAGGAGCTTGAAGAAATGCAAGTGAAGTGAAGTGAAGTGAAGTGA 2923
 QY 2556 TTGCAAAATTTAAGCTTAATTAACGATGATGATGATGATGATGATGATGATGATG 2615
 Db 2924 TTGCAAAATTTAAGCTTAATTAACGATGATGATGATGATGATGATGATGATGATG 2983
 QY 2616 TGAAGCAATAGCA-----GTGCGCTTCTTGAAGAGGAAACAAAGCCACATG 2669
 Db 2984 TGAAGCAATAGCA-----GTGCGCTTCTTGAAGAGGAAACAAAGCCACATG 3043
 QY 2670 GCGAAGACGCTTTCAACATTTGACAGTAAATGTAATGTAATGTAATGTAATGTAAT 2729
 Db 3044 GCGAAGACGCTTTCAACATTTGACAGTAAATGTAATGTAATGTAATGTAATGTAAT 3103
 QY 2730 CAATTTACTTCTTTATTTGTGCTAATTAAGGAGTAAATTAATTAATTAATTAATTA 2789
 Db 3104 CAATTTACTTCTTTATTTGTGCTAATTAAGGAGTAAATTAATTAATTAATTAATTA 3163
 QY 2790 GAGGAGATTAATTAATTTGCTGTGCAACAGGCAAGGCAAGGCAAGGCAAGGCAAG 2849
 Db 3164 GAGGAGATTAATTAATTTGCTGTGCAACAGGCAAGGCAAGGCAAGGCAAGGCAAG 3223
 QY 2850 TTAATCTTGTGTAAGCAAGATTAATCAACCTTATGATTAAGTCAATTTACTTTA 2909
 Db 3224 TTAATCTTGTGTAAGCAAGATTAATCAACCTTATGATTAAGTCAATTTACTTTA 3283
 QY 2910 GAAATATGCAACGTTGATGAGGATGATGATGATGATGATGATGATGATGATGATG 2969

81 CTTATTTTAAACCGCTTCATTCATTAAGGATATGATGCAAGCGTGGCTGTGAC 140
73 CTTCTTACTAGCCATATGCTGTGTTCCGATTTCTCCAGACTTGGCGGAC 132
141 ACTATTTTGGATTTGATTAACAAATATATCGTATTTTCCGAGATTAAGGAACTTC 200
133 ACTATTTTGGATTAACAAATATATCGTATTTTCCGAGATTAAGGAACTTC 192
201 ACAGTTGGGCTCAAAATATTAAGTTTAAACAAAGGCAATTAAGTTGGACATCA 260
193 GCACTGGGCGCAAAAGATTTAGGTTTAAACAAAGGCAATTAAGTTGGACATCA 252
261 ATGCAAAAGCCCGATGATTTTCTGTAGTTCAGTAAACGAGCGTGGCACTTG 320
253 ATGCAAAAGCCCGATGATTTTCTGTAGTTCAGTAAACGAGCGTGGCACTTG 312
321 GTTGAATATCATATTTTGGAGCGTGGCACTTAAGGATATTAAGATTTT 380
313 GTGGGCGATCAATATTTTGGAGCGTGGCACTTAAGGATATTAAGATTTT 372
381 GTTGAAGGAGCAACCCGATTCATCATCTTTTACTTATTAAGTTGTAACAAAT 440
373 GGTGGAGAGAGAGAAATCCCGATTCAGACCTTTTCTTCAAAATTTGTAAGAAAT 432
441 AACTCAAAAAGATATTTTCACTTATGAGAGCAATTAACATTAATCAATCAAT 500
433 AATTAATAGCTGCAATTCACCTTTCAACGCGGATTAACATTAATCAATCAAT 492
501 AATTTGTTAGAGAGGCTCCATTTGATTAAGGATTTGATTAAGGATTTGAT 560
493 AATTTGTTAGAGAGGCTCCATTTGATTAAGGATTTGATTAAGGATTTGAT 552
561 TCAGATAGCAAAATATCCAGACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 620
553 TCCGATTAAGGAAATATCCGAGCTGTGCGCATCGGCTGAGAGACCACTATTTGCGT 612
621 AATGATCAAGCAAAAGGCAACGTTGCGGCTGATTAATTTGATTAAGGATTTGAT 680
613 TATGATGATGACAAACAGCGGATTTATCTTATCCGAGCATTTGATTTGAGGCAAT 672
681 AACAATATGCGGAGAGGATTAAGGATTTGATTTGATTTGATTTGATTTGATTT 740
673 AACAATATGCGGAGAGGATTTGATTTGATTTGATTTGATTTGATTTGATTT 732
741 GCGGAGATATGTTGATTAACGATTTGATTTGATTTGATTTGATTTGATTTGAT 800
733 GCGAGATATGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 792
801 TTTATTTATGATGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 860
793 TTTATTTATGATGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 852
861 CTTTGAAGGAAAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 917
853 CTTTGAAGGAAAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 912
918 AATTTGAAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 977
913 AATTTGAAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 972
978 AATTTGAAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1037
973 TTTATTTATGATGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1032
1038 AATTTGAAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1088
1033 CCAAGCTTTAAAGTACAGATTCGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1092
1089 GTTCAATATCTAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1148
1093 GAACGATTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1152
1149 AGCTATATTTTATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1208

1153 AACCTTCTTTATGATTAACGCAACGCAAACTCATCTTATTAACAAATCAACCA 1212
1209 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1268
1213 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1272
1269 TGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1328
1273 TGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1332
1329 GTGGAACATGATGATCTTTCTAAATTTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAG 1388
1333 GTGGAACATGATGATCTTTCTAAATTTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAG 1392
1389 GAAATTAAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1448
1393 GAAATTAAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1452
1449 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1508
1453 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1512
1509 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1568
1513 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1572
1569 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1628
1573 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1632
1629 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1688
1633 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1692
1689 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1748
1693 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1752
1749 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1808
1753 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1812
1809 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1868
1813 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1872
1869 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1928
1873 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1932
1929 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1988
1933 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1992
1989 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2048
1993 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2052
2049 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2108
2053 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2112
2109 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2168
2113 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2172
2169 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2228
2173 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2232
2229 GATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2288

Db 2233 ACTAAGACGACATCAGCGGCAATGTCAGCTTGGCCATCAGCTCATTTAAATCTCA 2292
 QY 2283 GGTTCAGCAAACTTAATGSCATGTCATTTAACTAACTACAGCCATTTACATAGC 2348
 Db 2293 GGGCTTCACACTCACTACGCGCAATCTTAGTCAAATGCGCATACGTTATACAGTCAGC 2352
 QY 2349 AACATCCACCCCAATAGGCAATTTTCG----- 2377
 Db 2353 CACACGCCACCCCAAAACGCGCACTTAGCTCGTGGCAATGCCAGCAACATTTAAT 2412
 QY 2378 ----- 2377
 Db 2413 CAAGCCATTTAAACGGCAACATCGGCTTCGGCAATGCTTCATTTAATCTAAGCAAC 2472
 QY 2378 ----- ACTTCCGCAATTCACATGSCAGGTCGATAT 2411
 Db 2473 AACCCGCAAAAAGGCACTGACGCTTTCGCAACAGCTTAAGCAAAAGTCAGCAT 2532
 QY 2412 GCAATCTGACGTAATGTCATTTACGATTCAGCTCAATTTTCTTTAAAAACAGC 2471
 Db 2533 TCACCATCAGCGCAATGTCTCCCTAGCCGTAAGGCAATTCATTTGAAAAACAG 2592
 QY 2472 CATTTTGGACCAATTCAGGAGCAAGGCAACAGTCAGCTTGAGAAATGCGACT 2531
 Db 2593 CGCTTACCGCAACTCAGCGAGCAAGATACAGCTTACATTTAAAGACAGCGCA 2652
 QY 2532 TGACATGCTAGCGATACATTTGAGATTTAAAGCTAAATACAGTACATCAGC 2591
 Db 2653 TGACCGCTGCGCTCAGGACAGCAATTTAGCAATTTAACTTACAGCCACATTTAC 2712
 QY 2592 TTAAATCAGCTTAT-----TCAGTACCTCAACATATAGC 2627
 Db 2713 CTCAATTCGCTTATCGCCAGATGCTGAGGCGCGCAACCGGAGTGTGTGACAGAG 2772
 QY 2628 CCGCTGCGCTTAT-----TAGAGCGGAAACAGCCACATCGCGCAACAT 2678
 Db 2773 CCGGCGCGCTTATCGCGCTTATCGCTTATCGCTTATCGCTTATCGCTTATCG 2832
 QY 2679 CGTTTCAACATTCAGTAAATGTAATGATGGGCAAGGCACTTCATTCATTTACT 2738
 Db 2833 CGTTTCAACGCTCAGGTAAGGCAATTTGAAAGGCAATTCGCTTATG 2892
 QY 2739 TCATCTTTATTTGCTTAAAGGCAATTAATTAATTAATTAATTAATTAATTAATTA 2798
 Db 2893 TCGGAATCTTTCGCTTACGAGGCAATTAATTAATTAATTAATTAATTAATTAAT 2952
 QY 2799 TACATTAATCTGTCGACACAGGCAAGGCAAGCCGAAACCTTGAAGATTTAATCTTG 2858
 Db 2953 TACCTTGGCGGTCACACATTAACGCGAAGGCAAGCCGTAAGCTTGAATTAATTAAG 3012
 QY 2859 GTTGAAGCAAGATTAATCAACGTTATCAGATAGCTCAATTTTCTTTGAAATGAC 2918
 Db 3013 GTTGAAGGGAAGGCAACAAACCGCTTCCGAAACCTTAATTTTACCTTGAAGAA 3072
 QY 2919 CAGCTTATGTCAGGTCATTTAGTTAATTTAGTAAAGATGATGCGAATTCGCTTG 2978
 Db 3073 CAGCTGATGTCGCGCTGCTGCTTACCAATTCATTCGCAAAACGCGCAAGTTCGCTG 3132
 QY 2979 CATACCCCAATTAAGGCAAGGCAATTCAGATGATTTAGTAAAGGCAAGGCAAG 3038
 Db 3133 CATATTCGCTCAAAAGGCAAGGCTTTCGCAACCTCGGCAAGGCAAGGCAAGGCA 3192
 QY 3039 CGAATCTAGAACCA-----ACAACTTGAACGATGCTTAAACACAAACAGTGAACCA 3095
 Db 3193 CAGGCGGAAAAAGCAACGCGCAACGCTTACGCGCTGATTCGCGCGCGGAGATGCC 3252
 QY 3096 AAGTGCCTGCAAGAGGCAAGGCAAGGCTTTCCTATACCTTGCCTGATCAACAG 3155
 Db 3253 GCGGAAAGCAAGAGGCTTTCGCAACGCGCGCGCGCAAGGCAAGGCAAGGCAAG 3312
 QY 3156 CTGTTAAAGCATTTAGAGCAAAACAGCTGATGATCTGAAACCAAAAAAGTAG 3215
 Db 3313 ATATATGACGCGAGGAAAGGAAAAACGCGTGAAGCGGAT----- 3354

QY 3216 GCAAAAACAAAAAGTGGCGTCAAAAAAGCAGTGTTCATTCCTTCCTTATCA 3275
 Db 3355 -----AAGACAGGCTTTCGCAACAGCGCAAGGCAAGGCAAGGCAAGGCAAG 3406
 QY 3276 AGCTGTTTCGCTTACAGCCGCACTTGAAGTTATGATGCGCAACAGCAATGCAAAAA 3335
 Db 3407 CGGCTTTCCTCCCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3459
 QY 3336 GATCGTTCGCTCAAGAGAGAGCGGAAAAACAGCAACAAAAAGCTTATGACGCT 3395
 Db 3460 CAACCGCAACTCAACCCCAACG-----CAACCGCACTGATCAACCGCT 3504
 QY 3396 TATTCATTAATGCTGATCAGATTTATTCGAAACGTAATTAATTAATTAATTAATTA 3455
 Db 3505 TATTCATTAATGCTGATCAGATTTATTCGCAACGCTTACAGAGGCTTTCGCGTACG 3564
 QY 3456 GATGATTAATGCTGCTTTTTCGATCAGCAACATTCGCGTGTGCAAAATATCGCA 3515
 Db 3565 GACGATTTGACCGCGCTGTTTCGCAAGACCGCGCGCAACGCGTTTTCGCAACGCGCATC 3624
 QY 3516 CAGGATTAAGAGGCTATGATTTGATGCTTTCGCTTATCAGCAGCAAAAGCAAC 3575
 Db 3625 CGGACACCAACATCACCGTTCGCAAGATTTTCGCGCTTACCGCAAC---AACCGAC 3681
 QY 3576 TTAAGTCAATTTGAGGAGCAAAACCTTAGCTAATGAGCAATTTGAGGCAAGTTTCG 3635
 Db 3682 CTGCGCAATTCGCTTATGCAAGAAACCTGCGCAAGGCGCGCTGCGCATCTGTTTCG 3741
 QY 3636 CATACCGCTTCAGATTAATTCCTTATGATGCAAGGTTAAATACAGGATTTACGATG 3695
 Db 3742 CACACCGGACCGAAACACCTTCGCAAGACGCGATCGGCACTCGGCAAGGCTTCGCAAC 3801
 QY 3696 ATGTGCGTTTTCGCAATTCATTCATGCGCGATTTCAATTTGTTGTAACGTTGCAACG 3755
 Db 3802 GCGCGCGTTTTCGCGCAATTCAGCAATCGCGAGGTTGATTCGCAATTCAGCAAGCGCGCG 3861
 QY 3756 GGAATCAGTGGAGTAAATGCTGAAAGCAAGCCGAAATTCATTCGAAAGCGATA 3815
 Db 3862 GGTTCAGACCGCGAGCTTTCAGACGCGATCGAGGCAAAATTCGCGCGCTGCTG 3921
 QY 3816 AATTAAGCGTGAATCAAGTATTCAGTTCGTTTGAAGCAATTTGAGCATTCAGCTTAT 3875
 Db 3922 CATACCGCATTCAGGCAAGTACCGCGCGCTTTCGCGGATTCGCAATTCAGCAACCTGAC 3981
 QY 3876 TTTGAGTAAATGCTATTTTATGACGTAATTAATTCATTCAGTGAAGGAGAGTG 3935
 Db 3982 ATCGGCGCAAGGCTATTTTCGCAAAAGCGATTAACGCTACGAAACGTCATATTC 4041
 QY 3936 AAAAGCGTACCTTCATTTATGCTAATTAATGCTGATTCAGTGAATTCATTT 3995
 Db 4042 GCGACCGCGCTGCTGCTTACCGCTACCGCGCGGATTTAGGCAAGTATTCATTC 4101
 QY 3996 ACTCGACAGTAAATTCAGGCTTACGCTTATTCGCTTATTCGCTTATTCGCTTATTC 4055
 Db 4102 AAACCGGCGCAACATTTTCATTCAGCTTATTTAGGCTGCTTATTCAGTACCGCT 4161
 QY 4056 AAGCTTAAGTAAACAGCGTAAATTCAGGCTTTCGCAACAGCAATTTGAGCGTTAT 4115
 Db 4162 TCGGCAAGGCTGCAAGCGCTCAATTCGCGCTTATTCGCTTATTCGCGCAAAAC 4221
 QY 4116 TGGCAAAAGAGTGGGATTAAGGCAAGAAATTTTCAATTTTCGCTTATTC 4175
 Db 4222 CGGATGCGGATGAGGCGCTTAAGCGCGAATTCAGGCTTTCAGCTGCTCCCTCAACGCT 4281
 QY 4176 TCAATATTCAGGTTCAATTCGCGCAACAGCAAAATTCGCGGCAATTTGAGGCTAT 4235
 Db 4282 GCGCGCGCAAGGCGCGCACTGCAAGGCGCAACAGCGCGGCTCAATTAATTCGCTAC 4341
 QY 4236 CGTGTGTA 4244
 Db 4342 CGCTGTTA 4350

RESULT 12

US-10-282-28843

Sequence 28843, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forebych, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

PRIORITY FILING DATE: 2003-02-20

PRIORITY APPLICATION NUMBER: 60/191,078

PRIORITY FILING DATE: 2000-03-21

PRIORITY APPLICATION NUMBER: 60/206,848

PRIORITY FILING DATE: 2000-05-23

PRIORITY APPLICATION NUMBER: 60/207,727

PRIORITY FILING DATE: 2000-05-26

PRIORITY APPLICATION NUMBER: 60/230,335

PRIORITY FILING DATE: 2000-09-06

PRIORITY APPLICATION NUMBER: 60/230,347

PRIORITY FILING DATE: 2000-09-09

PRIORITY APPLICATION NUMBER: 60/242,578

PRIORITY FILING DATE: 2000-10-23

PRIORITY APPLICATION NUMBER: 60/253,625

PRIORITY FILING DATE: 2000-11-27

PRIORITY APPLICATION NUMBER: 60/257,931

PRIORITY FILING DATE: 2000-12-22

PRIORITY APPLICATION NUMBER: 60/267,636

PRIORITY FILING DATE: 2001-02-09

PRIORITY APPLICATION NUMBER: 60/269,308

PRIORITY FILING DATE: 2001-02-16

Remaining Prior Application data removed - See file wrapper or PAM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 28843

LENGTH: 4404

TYPE: DNA

ORGANISM: Neisseria gonorrhoeae

US-10-282-28843

Query Match 30.0%; Score 1294.8; DB 13; Length 4404;

Best Local Similarity 59.1%; Pred. No. 2,1e-294;

Matches 2568; Conservative 0; Mismatches 1567; Indels 213; Gaps 12;

Qy 86 TTTTACCGCTGCATTTGATAGGATAGTATCGACGCTGCTGTCACACTTA 145

Db 78 TTACTAGCCATATGCTGTCGTCGTCATTTGCCCCAGCCCGGGCGACACACTTA 137

Qy 146 TTTTGGATTGATTAACAATATATCGTATTTTCCGAGATTAAGGAGAGTTCAACT 205

Db 138 TTTTCGATCACTACCAATACATCATCGGACCTTCCGAAAATAAGGCAATTTGCACT 197

Qy 206 TGGGCTCAAAATATTAGGTTTATCAAAAGAGGCAATTTGGGCACTCAATGAC 265

Db 198 CGGGCGAAGATATTAGGTTTACCAAAAAGGAGGAGTTGCTCGCAATTCATGAC 257

Qy 266 AAAAGCCCGATGATGATTTTCTGTAGTTCACGTAAACGCGCTGGACGCTTGCTGA 325

Db 256 GAAAGCCCGATGATGATTTTCTGTAGTTCACGTAAACGCGCTGGACGCTTGCTGA 317

Qy 326 AATTCATATATTGAGCGCTGGACATTAAGTATATCAAGATTTGATTTTGGTGC 385

Db 318 CGATCAATATATTGAGCGTGGCAATTAACGCGCTATTAACATGTTGATTTGGTGC 377

Qy 386 AGAGGAAACACCCGATCAATCGTTTACTTATAGATTTGAAGAAATACATA 445

Db 378 GAGGAGAACATCCGATACACCGCTTTTCTTCCAAATTTGAAAAGAAATATTA 437

Qy 446 CAAAAG---ATATTTACATCTTATGAGAGCATTAATCCATATCCAGATTACATA 502

Db 438 TAAAGAGGACTAAGGCGCATCTTATAGGCGGATATCAATAGCGGCTTGACATA 497

Qy 503 ATTGTTACAGAGCGGCTCCATTTGATATGACTTGAACATGAATGCGCTTATTC 562

Db 498 ATTGTCACATGCGAAGCTGTTGATGATGACAGCTTATGATGAGTGGAAATACGC 557

Qy 563 AGATGAAACAAATATCCAGACGTTGATGAGCTGGAGCGGAGTTTGGGAAA 622

Db 558 TGAATTAATTAATACCTGATGATGATGATGAGGAGGAGACAAATTTGGCGTC 617

Qy 623 TGATCAGACAAAGC-----GACCAAGTTCGGGTGATATCATTA 664

Db 618 TGATGAGAGCAACCCATTAACCGGAAAGTCAATATGCAATGCAAGGATATCTTG 677

Qy 665 TCTGACAGCTGGCATACACAAATCAGCGTGGAGAGGATATGATTTGATTTGGG 724

Db 678 GCTCGTGGTGGCAATACCTTTGCAAAATGATGAGTGGTGGACAGTCAACTTAAG 737

Qy 725 AGCGATGTTGTTAAACGGGAGAAATATGTCATTAACGATGGAGGCTCAAGGGAGA 784

Db 738 TACGCAAAATTAATTAACATAGCCATATGTTTATCCAAAGGAGGCTATTTGGCA 797

Qy 785 CAGTGGTCTCCGATGTTATTTATGATGCTGAAAAACAAATGTTATTAATGAGAT 844

Db 798 CAGTGGCTCACCAATGTTATGATGATGATGATGATGATGATGATGATGATGATG 857

Qy 845 ATTAGGGAAGGCAACCTTTGAGGAAAGAAATGTTGATTTGATTTGTTGCAATTC 904

Db 858 ATTGCAAAAGGCAACCTTTGAGGAAAGAAATGTTGATTTGATTTGTTGCAATTC 917

Qy 905 TTAATTT---TGATGAAATTTTGAAGAGATTTATACATCATCTTACACCGAGCTGG 962

Db 918 TTGTTCTATATTAATTTCTTCTGAGATGATGATGATGATGATGATGATGATGATG 977

Qy 962 TAATGAGTGAACATTTGAGGAAATGATATGTCAGGGGTCTAT-----AAGT 1013

Db 978 AATGGGAAATATCTTTTAAAGCAATTAATGCGGAGGAAATATGATGCAAAAGA 1037

Qy 1014 CAGAAATCAGAAATACATCAGAAATTAATACG-----TAGCAATATAGAG 1063

Db 1038 TAACACTATTTCTGACCTTATATGATTAATAAACACGAAACGTTCAATTTGTTATGTTTC 1097

Qy 1064 TTACCTTTGAAGAGAGGATTAAGTTCATATCTTATGATGAGGAGGAGGAGGAGGAGG 1123

Db 1098 TTTATCCAGACAGAAAGAAACCTGTTTATCATGCTGCA---GGTGGGCTCAACGTTA 1154

Qy 1124 TTCTCAGCTTAAACATGAGAAAGGCTATATTTTATGATGATCAAAACAGAGATCAT 1183

Db 1155 TCAGCCAGACCTGATATATGAGAAATATTTCTTATGCAAAAGGAAAGGAGTAAAT 1214

Qy 1184 AATCTTGGATCTACATTTAACAGGGGGGGGCTTTTATTTGAGGTAATTTTAC 1243

Db 1215 GATACCTAACGCAATCAACCAAGGGGGGGGCTTTTATTTGAGGTAATTTTAC 1274

Qy 1244 AGTATCCAAATTTCAACCAACTTGGCAAGAGCTGTCATATGTAAGTGAATAATAG 1303

Db 1275 GGTCTGCTTAAACAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1334

Qy 1304 CACGTTACTTGAAGATTAATGCGTGAACATGATGATCTTTCTAATATGTTAAAG 1363

Db 1335 TACCGTTACTTGAAGATTAATGCGTGAACATGATGATCTTTCTAATATGTTAAAG 1394

Qy 1364 AAGCTGACCTTGAACCAAGGAGAAATTAAGGTTGATCAGCGAGGAGGAGGATTA 1423

Db 1395 CAGCTGCTGTTCAACCCAAAGGAGGAAACCAAGGCTGCTGATGAGGAGGAGGAGGAGG 1454


```

Db      3600  CAAAGGCTTTTGGCCGTAACGACGATTTGACCGCTGTTGCCGAAAGCCGCCGAA 3659
Qy      3494  TGCCGTGTGACAAATATGCAACGATMAAAAGCTATGATTTGATGCTTCCTG 3553
Db      3660  CGCGTTTGGACAGGGGCAATCCGGGACACCAACACTACGTTCCGAGATTTCCGGC 3719
Qy      3554  TTATACAGACGAAAGAACTTACCTGAAATGGGCTGCAAAAAGCTTACTATGG 3613
Db      3720  CTACCGCCAAAC--AAACCGACCTGCCCAATTCGATGTGAGAAAACCTCGAGAGGG 3776
Qy      3614  ACGAATGGGGGAGTTTTCTGCAAGCGCTTCAGATTAATACCTTTGATGAAAGGTTAA 3673
Db      3777  GCGGCTGGCATCTGTTTGTGCAACGACGACCGGAAACCTTTGACGAGCGCATGG 3836
Qy      3674  AAATACGACGACATTAAGATGATGTCGGGTTTTCGCCATATCAATGGGGCGATTACA 3733
Db      3897  CAATCGGACACGGCTTGCCACGCGTGGCGTTTTCGGGCAATAGGCGCATCGAGTTGCA 3896
Qy      3734  ATTGTGTAAACGTGGGAAACGGGATCGTGGGATTAATAGGCTGAAAGAACAGCCG 3793
Db      3897  CATGGCATCGAGGGGGGGGGGTTTGTGTAGCGGACGCTTTCAGACGGCATCGAGG 3956
Qy      3794  AAAATATTCGAGAAAAGGATTAATTAATGCGGTAATGCAAGTTATCAGTTTACG 3853
Db      3957  CAAATCCGCCGCCCGGCTGCTGCAATTCGCGCATTAAGCAAGATACCGCGCAGTTTCG 4016
Qy      3854  GCAATTTGGCATTCAGCCTTAATTTTGAAGTTAATGCTATTTTAAAGTGAATTA 3913
Db      4017  CGATTCGGCATCGAACCGGACATCGGCGGCAACGGGCTATTTGCTCAAAAAGCGATTA 4076
Qy      3914  TCAATCTGAGAGAGTGAAGAGTGAAGAAAGCTTACCTTGCATTTATTCGTTATGCTGG 3973
Db      4077  CCGATACGAAAGCTGCAATATGCGACCCCGGCGCTTTCATTAACCGCTACCGCGGG 4136
Qy      3974  CATTGAGTTGATTAATTAATTAATTCGCAAGTAAATAGCGTTAAAGCTTATTTCT 4033
Db      4137  CATTAAGCAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4196
Qy      4034  CGTCATTTATGTTATGTTTGAAGCTTAACGTAACGTAACCAACGTAATCTACGGTGT 4093
Db      4197  CCGTCTTATACCGATGCGGCTTCGGGCAAGGCGAAGCGGCTGATTAACCGCTATTT 4256
Qy      4094  GCAACAAACATTTGACGTTATTTGGCAAAAAGAGTGGGATTAAGCGCAAGATTTTCA 4153
Db      4257  GCGCGAGATTTTGGCAAAACCGCAGTGGGATTTGGGCGTAAACCGCAATCAAGG 4316
Qy      4154  TTTCGAATTTTCGCTTTTATCTCAAAATCTCAAGTTTCACTCGGCAACAGCAAAA 4213
Db      4317  TTTCAGCTGTCTTCCACAGCTGCGCGCCGCAAGGGGCGCAATTTGAAGCGCAGCAG 4376
Qy      4214  TGTGGCGGTGAATTTGGGCTATTCGTTGG 4241
Db      4377  GCGGGGATCAATTTAGGCTACCGCTGG 4404

```

RESULT 13

US-10-282-122A-22193
 : Sequence 22193, Application US/10282122A
 : Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zykkind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

```

: TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: ELITRA.034A
: CURRENT APPLICATION NUMBER: US/10/282,122A
: CURRENT FILING DATE: 2003-02-23
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,048
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: Remaining Prior Application data removed - See File Wrapper or PAM.
: NUMBER OF SEQ ID NOS: 78614
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 22193
: LENGTH: 5085
: TYPE: DNA
: ORGANISM: Haemophilus influenzae
US-10-282-122A-22193

Query Match      4.7%; Score 202; DB 13; Length 5085;
Best Local Similarity 60.9%; Pred. No. 1,76-36;
Matches 347; Conservative 0; Mismatches 220; Indels 3; Gaps 1;

Qy      1176  GATCATTAATCTTCGATGTGATTAACCAAGGCGCGTGTCTTATTTGAGGCT 1235
Db      1183  GGAACCTTACCTTAAATATATATGATCAACGCGCGGCTTGTCTTGAAGGA 1242
Qy      1236  AATTTAGATGATTCGAATTC--TAACCAACTTGGCAAGAGTGGCATCATGTA 1292
Db      1243  GATTTAGAGTTAAAGGACTTGTGATGATGACCTTTGAAAGAGCTGGGTTCTGTT 1302
Qy      1293  AGTGAATAGACCGTTACTTGGAAAGTAAATGGCGTGAACATGATGACTTTCTAAA 1352
Db      1303  GGTGATGAAAAACAGTAAGTGAAGAAAGTACATACCCGAATCTGATGTTAGCTAAA 1362
Qy      1353  ATTGTAAAGAAATGTGACGTTCAAGCCAAAGGGGAAATTAAGTTCATAGCTTA 1412
Db      1363  ATGGGCAAGAAACATTAATTTAGAGAAAGGAAAGAAATTAAGTTGCTTAAAGTG 1422
Qy      1413  GGGGATGTAAGATCTTTTGGAGCAGGACAGACGCAAGGCAACAAACAGCCTTT 1472
Db      1423  GCGCAATGATCTTATCTTAAACAAACAGCTATCCATTAATTAAGTTAAAGCCTTT 1482
Qy      1473  AGTGAATGGCTGTGTTAGCGGACAGGACCTTTCAATTAAGATGAATCAATTT 1532
Db      1483  TCACAGTAGTAATAGTAAAGTGTGCTCACTTTGTAATTAAGATGAACAGTA 1542
Qy      1533  GATACGATTAATTTTATTTGGCTTGTGTGTGTGCTTAATCTTAAGGCGCATTC 1592
Db      1543  GATCCAAATCCCATTTTCTTTGGCTTAAAGGTGTGATTAAGCCAAATGCAATAT 1602
Qy      1593  TTAACTTTAAAGTATCCAAATACGACGAGGGGCAATGATTTGACCATATATCA 1652
Db      1603  CTCACCTTTGAACATATCCGTAATATTAATGATGGCCAGAGTAAATACAAATAC 1662
Qy      1653  ACTGAAGCGCTAATGCTAATTAATCTGGGAAAGAAAGCAATTTCTTCAATTAAG 1712
Db      1663  AGCAAAACCTTAATGTAATTAATCTGGGAAAGTCTAATTAAGATCCAAATACAT 1722

```

Qy 1713 AATATTATATAAAGCTGATTACAGAAAGAA 1742
Db 1723 ACTCATATATAATATAGCGACAGATGAA 1752

RESULT 14

US-10-282-122A-29633
; Sequence 29633, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282.122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 29633

LENGTH: 5322

TYPE: DNA

ORGANISM: Neisseria meningitidis

US-10-282-122A-29633

Query Match 3.8%; Score 164.4; DB 13; Length 5322;
Best Local Similarity 55.1%; Pred. No. 1.3e-27;

Matches 368; Conservative 0; Mismatches 291; Indels 9; Gaps 2;

Qy 1156 ATTTATGATCAAAACAGATCATTAATCTTGATCGATTCGATTAACCAAGGCGCGG 1215
Db 1070 ATGTACCTTGAAACAGCATTTGGATTGGATCAAAACATCAACCAAGCGCGG 1129
Qy 1216 GTGCTTTATTTGAGGTAATTTTCAGT---ATTCGAAATTCACCAACTGCGC 1272
Db 1136 GGGGCTGTTTTCACAGCGATTCACAGTCAAGGTGCAAAATATGAGCATCTTGGC 1189
Qy 1273 AAGAGCTGCATATCATGTATGTAATATGACCGGTACTTGGAAAGTAAATGCGCTG 1332
Db 1190 TAGTGCAAGGATATGATGTGCGCAGCGCAAAATAGTGTGGCAAGTCAAAAATCCGA 1249
Qy 1333 AACATGATCACTTTTAAATATGTTAAAGAAATATGACGTTCAAGCCAAAGCGGAA 1392

Db 1250 ATGCGACAGATTTGCGCAAAATGCGCAAAAGCACATTGAAATTAACCGGACAGCGCTTA 1309
Qy 1393 ATTAAGTTGATCATAGGTAGCGATGTAAAGTATTTTGAGACAGGACGACAGCATC 1452
Db 1310 ACCAAGGCGATTAATTAAGTGGCGAGGTACGTTATTTGAAATCAACACCGATCCG 1369
Qy 1453 AAGGCAACAAACAGACCTTTGTAATTTGCTTGGTTAGCGGACAGAGCGGACTGTCAAT 1512
Db 1370 ACAAATAATCCAGGCTTTCTCCCAAGTGGGATTTGACAGCGGCGGCTACATGTGTC 1429
Qy 1513 TAAACATATATTAACATTTGATACGATATATTTTTCGCTTTGCTGCTGCT 1572
Db 1430 TGAACAGTCCGACCAATTAACCGAATATCTGTATTTTCGCTTTCCGTGCGCTGTT 1489
Qy 1573 TAGATCTTAAGGGGATTCATTAACCTTTAAAGTATCCAAATACGACAGAGGGGCAA 1632
Db 1490 TGAATGCCATGCAATGATCTTAATTTGAAACATGCGCAAGTGTATTAAGCGCGC 1549
Qy 1633 TGATGTGACCATTAATACATCAAGCGCTTAATGTCACTATTAATCAAGCAAGCA 1692
Db 1550 GATGTGCAACCAACAGACAGCGCGCTCGACATCAATTAACCGGCAATCATTA 1609
Qy 1693 TTGTTCTACTTAATGAAATTAATTTATTAATTTGATTCAGAAATGAAATTCCTTACA 1752
Db 1610 TTACCGCTCCT-----CAAAATCTTTGTTTATGAAATCCAAATGATTAACATGATG 1663
Qy 1753 ACGGTGTTTGGCGAAACGATTAATTAATTAACCAATGAGCGGATTAACCTTATTA 1812
Db 1664 ACGATTTATTTGCTATTAATTAACGCTATTAAGAAATCCATTCGCAAGCGAAATCTTAT 1723
Qy 1813 AACCAAC 1820
Db 1724 ACAAAC 1731

RESULT 15
US-10-282-122A-29078
; Sequence 29078, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282.122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

```
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29078
; LENGTH: 4779
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-29078
```

```
Query Match          3.6% Score 155.2; DB 13; Length 4779;
Best local Similarity 56.9%; Pred. No. 1.8e-25;
Matches 305; Conservative C; Mismatches 228; Indels 3; Gaps 1;
```

```
QY 1152 CTATATTTTATGATCAAAAACAGATCATTAATCTTGCATCTGACATTACCAAGG 1211
    |||||
Db 1399 CAATATGTTACTTTGAAGACACGGCACTTGTATTGATCAAAACATCAACCAAGGC 1158
    |||||

QY 1212 GCGGCTGCTCTTATTTGGGGTATTTTACGT--ATCTCAAAATTTACCAACT 1268
    |||||
Db 1159 GCGGCTGCTCTTATTTGAAAGGATTAACATGACGTAAGGCGCAATTAATGACATCACT 1218
    |||||

QY 1269 TGCAAGAGCTGCGATCACTGTATAGTAAATAGCACCGTTACTTGAAGTAATGCG 1328
    |||||
Db 1219 TGCTTAGTGCGGGATGTATGTGCGACGCGCAAAAAGTGTGCGAAGTCAAAAAT 1278
    |||||

QY 1329 GTGGAACATGATCGACTTCTTAAATTGGTAAGGACATTGACGTTCAAGCCAAAGG 1388
    |||||
Db 1279 CCGAATGCGGACGATGCGCAAAATCGGCAAGGCACTTGAATTAACGCGACAGGC 1338
    |||||

QY 1389 GAAATTAAGGTTCATAGCGTAGCGATGCTAAAGTCACTTTGAGAGAGAGCGAGAC 1448
    |||||
Db 1339 GTTACCAAGGCAATTAAAGTGGGAGCGGTATGCTGATCPAAAAAGCCGAT 1398
    |||||

QY 1449 GATCAAGGCAACAAACAGCCTTTAGTGAATTGGCTTGGTACGCGACAGGACTGTT 1508
    |||||
Db 1399 TCAAAACAAAAGTTCAAGGCTTCTCCCAAGTGGCATGTGACGGAGCGGATCAATTG 1458
    |||||

QY 1509 CAATTAACGATGATTAACATTTGATACGATTAATTTATTTGGCTTTCGTGTGCT 1568
    |||||
Db 1459 GTGCTGACAGTCCGACCCAGATTAAACCGAATATCTGTATTTCCGTTCCGTGGCGGT 1518
    |||||

QY 1569 CGCTTAGATCTTAAAGGCACTTCACTTAACCTTTAAAGTATCCAAATTAAGGAGAGG 1628
    |||||
Db 1519 CGTTGATGCCAATGGCAATGACTTTTGAACAATCCGCAAGCTGGATGAAGGC 1578
    |||||

QY 1629 GCAATGATTGTGAACCAATTAATACACTCAAGCGCTAATGTCACTATTACTGGGAA 1684
    |||||
Db 1579 GCGGCACTTGTCAACCAACACAGACACGCTTCCACATGACACTAACGGGTAA 1634
    |||||
```

Search completed: June 24, 2004, 20:24:32
Job time : 1730 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 07:38:46 ; Search time 13101 Seconds
(without alignments)
11599.790 Million cell updates/sec

Title: US-10-645-655-1

Perfect score: 4319

Sequence: 1 TCAATAGTGGTTACTAGT.....ACCTTTTATTCATATAT 4319

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3757330 segs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:*

```

1: /cgn2_6/ptodata/2/pna/US06US COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06 COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07 COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*

```

```

44: /cgn2_6/ptodata/2/pna/US100B COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US101A COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US101B COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US102A COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US102B COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US103A COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US103B COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US104A COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US104B COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US105A COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US105B COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US106A COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US106B COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US107A COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US107B COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US108A COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US108B COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US109A COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US109B COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US110A COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US110B COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US111A COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US111B COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US112A COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US112B COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US113A COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US113B COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US114A COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US114B COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US115A COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US115B COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US116A COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US116B COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US117A COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US117B COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US118A COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US118B COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US119A COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US119B COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US120A COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US120B COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US121A COMB.seq:*
86: /cgn2_6/ptodata/2/pna/US121B COMB.seq:*
87: /cgn2_6/ptodata/2/pna/US122A COMB.seq:*
88: /cgn2_6/ptodata/2/pna/US122B COMB.seq:*
89: /cgn2_6/ptodata/2/pna/US123A COMB.seq:*
90: /cgn2_6/ptodata/2/pna/US123B COMB.seq:*
91: /cgn2_6/ptodata/2/pna/US124A COMB.seq:*
92: /cgn2_6/ptodata/2/pna/US124B COMB.seq:*
93: /cgn2_6/ptodata/2/pna/US125A COMB.seq:*
94: /cgn2_6/ptodata/2/pna/US125B COMB.seq:*
95: /cgn2_6/ptodata/2/pna/US126A COMB.seq:*
96: /cgn2_6/ptodata/2/pna/US126B COMB.seq:*
97: /cgn2_6/ptodata/2/pna/US127A COMB.seq:*
98: /cgn2_6/ptodata/2/pna/US127B COMB.seq:*
99: /cgn2_6/ptodata/2/pna/US128A COMB.seq:*
100: /cgn2_6/ptodata/2/pna/US128B COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US129A COMB.seq:*
102: /cgn2_6/ptodata/2/pna/US129B COMB.seq:*
103: /cgn2_6/ptodata/2/pna/US130A COMB.seq:*
104: /cgn2_6/ptodata/2/pna/US130B COMB.seq:*
105: /cgn2_6/ptodata/2/pna/US131A COMB.seq:*
106: /cgn2_6/ptodata/2/pna/US131B COMB.seq:*
107: /cgn2_6/ptodata/2/pna/US132A COMB.seq:*
108: /cgn2_6/ptodata/2/pna/US132B COMB.seq:*
109: /cgn2_6/ptodata/2/pna/US133A COMB.seq:*
110: /cgn2_6/ptodata/2/pna/US133B COMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Match	100.0%; Score 4319; DB 1; Length 4319;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 4319; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Query 1	TCGATAGTCGTTTACTAGTATTTTAAATGAGAAAATTCTTAATTAATTAACATTA 60
Db 1	TCGATAGTCGTTTACTAGTATTTTAAATGAGAAAATTCTTAATTAATTAACATTA 60
Query 61	TGAAAAAACTGTAATTTTGGCTTAATTTTAAACCGCTGATTCATCAGGAGTACTTA 120
Db 61	TGAAAAAACTGTAATTTTGGCTTAATTTTAAACCGCTGATTCATCAGGAGTACTTA 120
Query 121	CGCAAGCGTGGGCTGTGTCACTTAATTTTGAGATTGATTAACCAATATATTCGTGATTTG 180
Db 121	CGCAAGCGTGGGCTGTGTCACTTAATTTTGAGATTGATTAACCAATATATTCGTGATTTG 180
Query 181	CCGAGATATTAAGGGAAGCTCAAGTTGGGGCTCAAAATATTAAGTTTAAACCAACAA 240
Db 181	CCGAGATATTAAGGGAAGCTCAAGTTGGGGCTCAAAATATTAAGTTTAAACCAACAA 240
Query 241	GGCATTTAGTTGGTGCATCATGACCAAAAGCCCCCAATGATTTTCTGTAGCTGCAC 300
Db 241	GGCATTTAGTTGGTGCATCATGACCAAAAGCCCCCAATGATTTTCTGTAGCTGCAC 300
Query 301	GTAAGCGGTTGGTGCAGCTTGTGAAATCATATATTGTGAGGCGACATAAGCTAG 360
Db 301	GTAAGCGGTTGGTGCAGCTTGTGAAATCATATATTGTGAGGCGACATAAGCTAG 360
Query 361	GATATACAGATGTGATTTTGGTGGAGAGGGAACAACCCCGATCAACATCGTTTACTT 420
Db 361	GATATACAGATGTGATTTTGGTGGAGAGGGAACAACCCCGATCAACATCGTTTACTT 420
Query 421	ATAAGATTGTATTAACCAATTAATCAAAAAAGATTAATTAATCGCTTGTGAGACAT 480
Db 421	ATAAGATTGTATTAACCAATTAATCAAAAAAGATTAATTAATCGCTTGTGAGACAT 480
Query 481	ACCATATTCACGATATACATTAATTCCTTACAGAGGCGGCTCAATATGATATGACCTTCA 540
Db 481	ACCATATTCACGATATACATTAATTCCTTACAGAGGCGGCTCAATATGATATGACCTTCA 540
Query 541	ATATGAAATGGACGATCTTATATGATATGAAACCAAAATATCCAGAACTGTGTGATCGGCT 600
Db 541	ATATGAAATGGACGATCTTATATGATATGAAACCAAAATATCCAGAACTGTGTGATCGGCT 600


```

Db 3121 GAGCAGGTTCCGATACCTGCTCATCAAGCCTGTTAAAGCATTTGAAGCCAAAC 3180
Qy 3181 AAGCTGAATGACTCTGTGAACACAAAAAGTAAAGCAAAAAACAAAAAGTGGCTGA 3240
Db 3181 AAGCTGAATGACTCTGTGAACACAAAAAGTAAAGCAAAAAAGTGGCTGA 3240
Qy 3241 AAAGCAGTGTGTTTCTGATCCCTGCTGATCAAGCCTGTTAAAGCATTTGAAGCCAAAC 3300
Db 3241 AAAGCAGTGTGTTTCTGATCCCTGCTGATCAAGCCTGTTAAAGCATTTGAAGCCAAAC 3300
Qy 3301 TTGAGTTATTTGATGCCCCACAGCAATCGAAAAAGATGCTAGCTAAAGAAAGCGG 3360
Db 3301 TTGAGTTATTTGATGCCCCACAGCAATCGAAAAAGATGCTAGCTAAAGAAAGCGG 3360
Qy 3361 AAAAACAAGCAAAACAAAAAGCTTGATGACCCGTTATCAAAAGTGGCTGTATGAT 3420
Db 3361 AAAAACAAGCAAAACAAAAAGCTTGATGACCCGTTATCAAAAGTGGCTGTATGAT 3420
Qy 3421 TATCTGACAGTAATAGTATGCTTCTGTTCAAGATGAAATGATGCTTTTGTAG 3480
Db 3421 TATCTGACAGTAATAGTATGCTTCTGTTCAAGATGAAATGATGCTTTTGTAG 3480
Qy 3481 ATCAAGCAATCTGCGCTGTGACAAATATGCAAGATGAAAGAGCTATGATCTG 3540
Db 3481 ATCAAGCAATCTGCGCTGTGACAAATATGCAAGATGAAAGAGCTATGATCTG 3540
Qy 3541 ATGCTTCCGCTGTTATGACAGCAAGAAACGAATGCTAAATGGGGGTGCAAAAG 3600
Db 3541 ATGCTTCCGCTGTTATGACAGCAAGAAACGAATGCTAAATGGGGGTGCAAAAG 3600
Qy 3601 CCTAGCTAATGAGCAATGGGGCAGTTTCTGCAATGCGTTCAAGTAATACCTTTG 3660
Db 3601 CCTAGCTAATGAGCAATGGGGCAGTTTCTGCAATGCGTTCAAGTAATACCTTTG 3660
Qy 3661 ATGAACAGTTTAAATTAACGCAATTAAGATGTCGGTTCCTTCCCATATTCAT 3720
Db 3661 ATGAACAGTTTAAATTAACGCAATTAAGATGTCGGTTCCTTCCCATATTCAT 3720
Qy 3721 GGGGGCAATTTAGCAATTTGGTGTAAACGTGGGAACGGGAATCAGTCCGATTAATG 3780
Db 3721 GGGGGCAATTTAGCAATTTGGTGTAAACGTGGGAACGGGAATCAGTCCGATTAATG 3780
Qy 3781 AAGAACAAAGCCGAAATTTATCGAAACGATTAATTTAGCGGTGAATGCAAGTTTC 3840
Db 3781 AAGAACAAAGCCGAAATTTATCGAAACGATTAATTTAGCGGTGAATGCAAGTTTC 3840
Qy 3841 AGTTCCGTTTGGGGCAATTTGGGCAATTTAGCTTATTTTGGAGTTATTCGTAATTT 3900
Db 3841 AGTTCCGTTTGGGGCAATTTGGGCAATTTAGCTTATTTTGGAGTTATTCGTAATTT 3900
Qy 3901 AACGTGAAATTTATCAATCTGAGAGAGTGAAGTGAAGACCTAGCCTTGAATTAATC 3960
Db 3901 AACGTGAAATTTATCAATCTGAGAGAGTGAAGTGAAGACCTAGCCTTGAATTAATC 3960
Qy 3961 GCTAATAATGCTGCAATGAGTTGATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
Db 3961 GCTAATAATGCTGCAATGAGTTGATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
Qy 4021 AGCTTATTTCTTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4080
Db 4021 AGCTTATTTCTTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4080
Qy 4081 ATCTCAGGTTGTCAGCAACACATTTGAGAGTTATTTGGCAAAAGAGTGGATTAAGG 4140
Db 4081 ATCTCAGGTTGTCAGCAACACATTTGAGAGTTATTTGGCAAAAGAGTGGATTAAGG 4140
Qy 4141 CAGAAATTTTACATTTCCAAATTTCCGCTTTATCTCAAAATTCAGAGTTCACACTG 4200
Db 4141 CAGAAATTTTACATTTCCAAATTTCCGCTTTATCTCAAAATTCAGAGTTCACACTG 4200
Qy 4201 GCAACGCAAAATGCGGCTGAATTTGGGCTATGCTGTTAAATTAATTAATTAATTT 4260

```

```

Db 4201 GCAACGCAAAATGTCGCGCTGAATTTGGGCTATCCTTGTAAATTAATTAATTT 4260
Qy 4261 ATGCTTATTTGTAAGCAAGTGGGTGAGATCAGATCCACCTTTCTTATTCATTAAT 4319
Db 4261 ATGCTTATTTGTAAGCAAGTGGGTGAGATCAGATCCACCTTTTATTCATTAAT 4319

RESULT 2
US-10-645-655-1
Sequence 1, Application US/10645655
GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Flehr, Hohnbach, Test, Albrighton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/645,655
FILING DATE: 20-Aug-2003
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-Aug-1994
ATTORNEY/AGENT INFORMATION:
NAME: Treccatlin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-5249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
FEATURES:
NAME/KEY: CDS
LOCATION: 60..4241
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-645-655-1

Query Match 100.0%; Score 4319; DB 53; Length 4319;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCATATGCTTTTAACTAGTATTTTAAATGCAAAATTAATTAATTAATTAATTAATTAAT 60
Db 1 TCATATGCTTTTAACTAGTATTTTAAATGCAAAATTAATTAATTAATTAATTAATTAAT 60
Qy 61 TGAATAAAAGCTATTTGCTTAAATTTTAAACGCTTGCATTTCAATGAGATAGTAT 120
Db 61 TGAATAAAAGCTATTTGCTTAAATTTTAAACGCTTGCATTTCAATGAGATAGTAT 120
Qy 121 CGAAGCGTGGCTGTCACACTTATTTGGGATTTGTAACAAATATATGATGATTTTG 180
Db 121 CGAAGCGTGGCTGTCACACTTATTTGGGATTTGTAACAAATATATGATGATTTTG 180
Qy 181 CGAAGAAATTAAGGAAAGTTACAGTTTGGGCTCAAAATATTAAGTTATTAACAAAG 240

```

Db 181 CCGAGATTAAGGAGAGTTCAAGTTGCGGCTCAAAATATATTAAGTTATTAACAAACAG 240
 Qy 241 GGCATTAAGTGGCAATCATTAACAAAGCCCGAGATGATTTTCTGAGAGTCAAC 300
 Db 241 GGCATTAAGTGGCAATCATTAACAAAGCCCGAGATGATTTTCTGAGAGTCAAC 300
 Qy 301 GTAAAGCGCTGGCAAGCTTGGTGAATATCAATATATTTGAGAGGCTGACATACATAG 360
 Db 301 GTAAAGCGCTGGCAAGCTTGGTGAATATCAATATATTTGAGAGGCTGACATACATAG 360
 Qy 361 GATATACAGATGTTGATTTTGTGTCAGAGGAAACACCCGATCAATCTGTTTACTT 420
 Db 361 GATATACAGATGTTGATTTTGTGTCAGAGGAAACACCCGATCAATCTGTTTACTT 420
 Qy 421 ATAGATTTGAAAAGAAATTAATTAAGAAATTAATTAATTAATTAATTAATTAATTAAT 480
 Db 421 ATAGATTTGAAAAGAAATTAATTAAGAAATTAATTAATTAATTAATTAATTAATTAAT 480
 Qy 481 ACCATTAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
 Db 481 ACCATTAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
 Qy 541 ATATGAATGGCAAGTACTTATTCAGATTAAGAAATTAATTAATTAATTAATTAATTAAT 600
 Db 541 ATATGAATGGCAAGTACTTATTCAGATTAAGAAATTAATTAATTAATTAATTAATTAAT 600
 Qy 601 CTGAGAGGCAAGTTTGGGAAATGATCAAGAAAGGCAACAAAGTTCGCGTGCATATC 660
 Db 601 CTGAGAGGCAAGTTTGGGAAATGATCAAGAAAGGCAACAAAGTTCGCGTGCATATC 660
 Qy 661 ATTATCTGACAGCTGGCAATTAACACATCAAGCTGAGAGCAAGTAAATTAATTAATTAAT 720
 Db 661 ATTATCTGACAGCTGGCAATTAACACATCAAGCTGAGAGCAAGTAAATTAATTAATTAAT 720
 Qy 721 TGGAGAGGCAAGTTTGGTGAATGAGAGAAATTAATTAATTAATTAATTAATTAATTAAT 780
 Db 721 TGGAGAGGCAAGTTTGGTGAATGAGAGAAATTAATTAATTAATTAATTAATTAATTAAT 780
 Qy 781 GGGACAGTGGTTCCTCGAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
 Db 781 GGGACAGTGGTTCCTCGAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
 Qy 841 GGATTAATCAAGGAGCAACCTTTGAAAGGCAAGAAATTAATTAATTAATTAATTAATTAAT 900
 Db 841 GGATTAATCAAGGAGCAACCTTTGAAAGGCAAGAAATTAATTAATTAATTAATTAATTAAT 900
 Qy 901 AATCTTAATTTGATGAATTTTGGAAAGATTTTCAATCAATCAATTTTCAACCGAGCTG 960
 Db 901 AATCTTAATTTGATGAATTTTGGAAAGATTTTCAATCAATCAATTTTCAACCGAGCTG 960
 Qy 961 AATCTTAATTTGATGAATTTTGGAAAGATTTTCAATCAATCAATTTTCAACCGAGCTG 960
 Db 961 AATCTTAATTTGATGAATTTTGGAAAGATTTTCAATCAATCAATTTTCAACCGAGCTG 960
 Qy 961 GTAAATGAGTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
 Db 961 GTAAATGAGTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
 Qy 1021 CAGGAATTAATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
 Db 1021 CAGGAATTAATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
 Qy 1081 AGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
 Db 1081 AGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
 Qy 1141 ATGAGAAAGAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
 Db 1141 ATGAGAAAGAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
 Qy 1201 TTAACAAAGGAGGAGGAGTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
 Db 1201 TTAACAAAGGAGGAGGAGTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
 Qy 1261 ACCAAATTTGGCAAGAGAGCTGGCATATGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1320
 Db 1261 ACCAAATTTGGCAAGAGAGCTGGCATATGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1320

Qy 1321 TAAATGCGTGGAAACATGATCGATTTCTAAATTTGGTAAAGAAACATGACGTTCAAG 1380
 Db 1321 TAAATGCGTGGAAACATGATCGATTTCTAAATTTGGTAAAGAAACATGACGTTCAAG 1380
 Qy 1381 CCAAAAGGAGAAATTAAGTGGATGATGACGTTGAGAGTAAAGTCAATTTTGGAGCAGC 1440
 Db 1381 CCAAAAGGAGAAATTAAGTGGATGATGACGTTGAGAGTAAAGTCAATTTTGGAGCAGC 1440
 Qy 1441 AGGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
 Db 1441 AGGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
 Qy 1501 GGCATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
 Db 1501 GGCATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
 Qy 1561 GTGCTGGTGGCTTATGATCTTAACGAGGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
 Db 1561 GTGCTGGTGGCTTATGATCTTAACGAGGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
 Qy 1621 AGGAGGAGGCAATGATTTGAGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
 Db 1621 AGGAGGAGGCAATGATTTGAGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
 Qy 1681 GGAACGAAGATTTGTTCTACCTTAATGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
 Db 1681 GGAACGAAGATTTGTTCTACCTTAATGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
 Qy 1741 AAATGCTCAACAGTGGTGGTGGGAAACAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
 Db 1741 AAATGCTCAACAGTGGTGGTGGGAAACAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
 Qy 1801 ACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
 Db 1801 ACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
 Qy 1861 TAAAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
 Db 1861 TAAAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
 Qy 1921 AGGCTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980
 Db 1921 AGGCTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980
 Qy 1981 TGGTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
 Db 1981 TGGTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
 Qy 2041 GCGGAAGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2100
 Db 2041 GCGGAAGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2100
 Qy 2101 ATATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
 Db 2101 ATATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
 Qy 2161 CAGATTTGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220
 Db 2161 CAGATTTGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220
 Qy 2221 ATTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
 Db 2221 ATTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
 Qy 2281 ATGTTAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2340
 Db 2281 ATGTTAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2340
 Qy 2341 CATTAAGCAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
 Db 2341 CATTAAGCAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400

2401 CCGTGAATATGCAAACTTGAACGGTATATGCAATTAACGGTACAGTCAATTTCT 2460
 2402 CCGTGAATATGCAAACTTGAACGGTATATGCAATTAACGGTACAGTCAATTTCT 2460
 2461 TAAAAAACAAGCAATTTTTCGACCAAAATTCAGGAGACAAAGGACAAACGATGCGTGG 2520
 2462 TAAAAAACAAGCAATTTTTCGACCAAAATTCAGGAGACAAAGGACAAACGATGCGTGG 2520
 2521 AAAATGCACTGGACAAATGCGTACGCACTATTCAGTGAATTTAAAGCTTAATACA 2580
 2522 AAAATGCACTGGACAAATGCGTACGCACTATTCAGTGAATTTAAAGCTTAATACA 2580
 2581 GTACGATCACTGTAATTAATGCTATTCAGTGAATTTAAAGCTTAATACA 2580
 2582 GTACGATCACTGTAATTAATGCTATTCAGTGAATTTAAAGCTTAATACA 2580
 2641 CATTAAGACGGAACAAACGCGCAATTCGCAACATTCGTTCAACATTCGCAATTA 2700
 2642 CATTAAGACGGAACAAACGCGCAATTCGCAACATTCGTTCAACATTCGCAATTA 2700
 2701 ATGTAATATGAGTGGGCAAGGCAATTCGCAATTTCTCAATTTTGGCTAATAAA 2760
 2702 ATGTAATATGAGTGGGCAAGGCAATTCGCAATTTCTCAATTTTGGCTAATAAA 2760
 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
 2762 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
 2821 CAGGCAAAAGAACCCGGAACCCCTTGACCAATTAATTTGTTGAAGCAAAAGATTAAT 2880
 2822 CAGGCAAAAGAACCCGGAACCCCTTGACCAATTAATTTGTTGAAGCAAAAGATTAAT 2880
 2881 CCGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2940
 2882 CCGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2940
 2941 GTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000
 2942 GTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000
 3001 AATTCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3060
 3002 AATTCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3060
 3061 TTGAACCACTGCTAAACCAACCAAGGTCGCAATTAATTAATTAATTAATTAATTA 3120
 3062 TTGAACCACTGCTAAACCAACCAAGGTCGCAATTAATTAATTAATTAATTAATTA 3120
 3121 AGACAGGCTTTCTGATACCTGCTGATCAAAAGCTTTAAAGCTTAATTAATTAATTA 3180
 3122 AGACAGGCTTTCTGATACCTGCTGATCAAAAGCTTTAAAGCTTTAAAGCTTTAA 3180
 3181 AAGCTGAATGATGCTGAACCAACCAAGGTCGCAATTAATTAATTAATTAATTAATTA 3240
 3182 AAGCTGAATGATGCTGAACCAACCAAGGTCGCAATTAATTAATTAATTAATTAATTA 3240
 3241 AAAAGAGCTTTCTGATACCTGCTGATCAAAAGCTTTAAAGCTTTAAAGCTTTAA 3300
 3242 AAAAGAGCTTTCTGATACCTGCTGATCAAAAGCTTTAAAGCTTTAAAGCTTTAA 3300
 3301 TTGAGGTTATGATGCTGAACCAACCAAGGTCGCAATTAATTAATTAATTAATTAATTA 3360
 3302 TTGAGGTTATGATGCTGAACCAACCAAGGTCGCAATTAATTAATTAATTAATTAATTA 3360
 3361 AAAAACAAGCAAAAGCAAAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 3420
 3362 AAAAACAAGCAAAAGCAAAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 3420
 3421 TATTCGCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3480
 3422 TATTCGCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3480
 3481 ATCAGCAATTCGCTGCTGATGCAAAATTAATTAATTAATTAATTAATTAATTAATTA 3540

3481 ATCAGCAATTCGCTGCTGATGCAAAATTAATTAATTAATTAATTAATTAATTAATTA 3540
 3541 ATGCTTCCTGCTTATTCAGCAACCAAAAGCAATTAATTAATTAATTAATTAATTAATTA 3600
 3542 ATGCTTCCTGCTTATTCAGCAACCAAAAGCAATTAATTAATTAATTAATTAATTAATTA 3600
 3601 CCGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660
 3602 CCGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660
 3661 ATGAACAGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3720
 3662 ATGAACAGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3720
 3721 GGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3780
 3722 GGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3780
 3781 AAGAACAAGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3840
 3782 AAGAACAAGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3840
 3841 AGTTCGTTAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3900
 3842 AGTTCGTTAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3900
 3901 AAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3960
 3902 AAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3960
 3961 GCTTAATGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4020
 3962 GCTTAATGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4020
 4021 AGCTTAATTTCTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4080
 4022 AGCTTAATTTCTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4080
 4081 ATTCAGGCTGTTGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4140
 4082 ATTCAGGCTGTTGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4140
 4141 CAGAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4200
 4142 CAGAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4200
 4201 GCAAAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4260
 4202 GCAAAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4260
 4261 ATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4320
 4262 ATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4320

RESULT 3
 US-10-687-046-1
 Sequence 1, Application US/10687046
 GENERAL INFORMATION:
 APPLICANT: St. Geme, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RT/DCF/DHR
 CURRENT APPLICATION NUMBER: US/10/687,046
 CURRENT FILING DATE: 2003-10-15
 PRIOR APPLICATION NUMBER: US/10/080,505
 PRIOR FILING DATE: 2002-02-22
 PRIOR APPLICATION NUMBER: US 08/296,791
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn version 3.1

SEQ ID NO 1
 LENGTH: 4319
 TYPE: DNA
 ORGANISM: Haemophilus influenzae
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (60)..(4241)
 OTHER INFORMATION:
 US-10-687-046-1

Query Match 100.0%; Score 4319; DB 53; Length 4319;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAATAGTCGTTTAACTAGTATTTTAAACGAAAAATTAATTATTAATAACATTA 60
 DB 1 TCAATAGTCGTTTAACTAGTATTTTAAACGAAAAATTAATTATTAATAACATTA 60
 QY 61 TGAATAAACTGTAATTCGTCTTAATTTTAAACGCTTGCAATTCATTAGGATAGTAT 120
 DB 61 TGAATAAACTGTAATTCGTCTTAATTTTAAACGCTTGCAATTCATTAGGATAGTAT 120
 QY 121 CGCAAGCTGGGCTGGTCAACTATTTTGGATTTGATTACCAATATATTCGATTTTG 180
 DB 121 CGCAAGCTGGGCTGGTCAACTATTTTGGATTTGATTACCAATATATTCGATTTTG 180
 QY 181 CCGAATAAAGGAAAGTTCAAGTTGGGCTCAAAATTTAGCTTTATTAACAAACAG 240
 DB 181 CCGAATAAAGGAAAGTTCAAGTTGGGCTCAAAATTTAGCTTTATTAACAAACAG 240
 QY 241 GGCATTAAGTTGTCACATCATGACAAAGCCCGCATGATGATTTTCTGATGTCAC 300
 DB 241 GGCATTAAGTTGTCACATCATGACAAAGCCCGCATGATGATTTTCTGATGTCAC 300
 QY 301 GTAAAGCTGGCAGCTGTTGTAATTAATTAATTTGAGGCTGGCACTAACATG 360
 DB 301 GTAAAGCTGGCAGCTGTTGTAATTAATTAATTTGAGGCTGGCACTAACATG 360
 QY 361 GATATACAGTGTATTTTGTGTCAGAGGAAACACCCGATCAACATGCTTTACTT 420
 DB 361 GATATACAGTGTATTTTGTGTCAGAGGAAACACCCGATCAACATGCTTTACTT 420
 QY 421 ATAGATTGTAAGGAAATTAATTAATTAATTTGATCTTATGAGGACATTT 480
 DB 421 ATAGATTGTAAGGAAATTAATTAATTAATTTGATCTTATGAGGACATTT 480
 QY 481 ACCATTAATCCAGATTAATTAATTTGTTACAGAGGCTGCAATTTGATGATCTG 540
 DB 481 ACCATTAATCCAGATTAATTAATTTGTTACAGAGGCTGCAATTTGATGATCTG 540
 QY 541 ATATGAATGGCAGTACTTATTCAGATTAAGAAACAAATATCCAGAGCTGTCTGATCG 600
 DB 541 ATATGAATGGCAGTACTTATTCAGATTAAGAAACAAATATCCAGAGCTGTCTGATCG 600
 QY 601 CTGAGAGGCGATTTTGGGAAATGATCAAGAAAGGACAAAGTTGCCGCTGATATC 660
 DB 601 CTGAGAGGCGATTTTGGGAAATGATCAAGAAAGGACAAAGTTGCCGCTGATATC 660
 QY 661 ATTATCTGACAGCTGGCAATACACATCACTGAGAGGATTAATGATTTGATTT 720
 DB 661 ATTATCTGACAGCTGGCAATACACATCACTGAGAGGATTAATGATTTGATTT 720
 QY 721 TGGAGAGCGATTTGTAAGCGGAGAAATATGTCATTACCGATTCAGGCTCAAGG 780
 DB 721 TGGAGAGCGATTTGTAAGCGGAGAAATATGTCATTACCGATTCAGGCTCAAGG 780
 QY 781 GGGACGTGTTCTCCAGATTTTATTAATGATGCTGAAAAACAAATATGTTATATG 840
 DB 781 GGGACGTGTTCTCCAGATTTTATTAATGATGCTGAAAAACAAATATGTTATATG 840
 QY 841 GGATATTACGGAGGCAACCTTTTGAAGGCAAGAAATGAGTTCAATTTGATGCA 900
 DB 841 GGATATTACGGAGGCAACCTTTTGAAGGCAAGAAATGAGTTCAATTTGATGCA 900

QY 901 AATCTATTATTTGATGAATTTTGAAGAGATTACATATCACTTACACCCGACTG 960
 DB 901 AATCTATTATTTGATGAATTTTGAAGAGATTACATATCACTTACACCCGACTG 960
 QY 961 GTATAGAGTGAACCAATTAATGAGAAATGATATGTCAGGGCTCTAATCAAGAAAT 1020
 DB 961 GTATAGAGTGAACCAATTAATGAGAAATGATATGTCAGGGCTCTAATCAAGAAAT 1020
 QY 1021 CAGGAATACCATCGAATTTAAATTAATGCTTGAATATGATTTACTTTGAAGAG 1080
 DB 1021 CAGGAATACCATCGAATTTAAATTAATGCTTGAATATGATTTACTTTGAAGAG 1080
 QY 1081 AGATTAAGTCAATATCTAGATATGAGAGCACTAATTTTATCTCAGCTTTAAAC 1140
 DB 1081 AGATTAAGTCAATATCTAGATATGAGAGCACTAATTTTATCTCAGCTTTAAAC 1140
 QY 1141 ATGAGAAACGCTATTTTATGATCAAAAAACAGATCATTAATCTTGCATCTGAC 1200
 DB 1141 ATGAGAAACGCTATTTTATGATCAAAAAACAGATCATTAATCTTGCATCTGAC 1200
 QY 1201 TTAAACAGAGGCGGGGTCTTATTTGAGGATTTTACATCTCCAAATCTTA 1260
 DB 1201 TTAAACAGAGGCGGGGTCTTATTTGAGGATTTTACATCTCCAAATCTTA 1260
 QY 1261 ACCAACTTGGCAGAGCTGGCATCATGTAAGTAAATATGACCCGTTACTTGAAG 1320
 DB 1261 ACCAACTTGGCAGAGCTGGCATCATGTAAGTAAATATGACCCGTTACTTGAAG 1320
 QY 1321 TAAATGCGTGGACATGATGCACTTTCAAAATTTGTTAAAGAAACATTGCACTTCA 1380
 DB 1321 TAAATGCGTGGACATGATGCACTTTCAAAATTTGTTAAAGAAACATTGCACTTCA 1380
 QY 1381 CCAAGGGGAAATTAAGTGTGATGAGGCAATGTAAGTAAATTTGAGAGG 1440
 DB 1381 CCAAGGGGAAATTAAGTGTGATGAGGCAATGTAAGTAAATTTGAGAGG 1440
 QY 1441 AGGACAGATCAAGGCAACAAAGCCTTTAGTGAATTTGCTTGTGTCGCGCAGAG 1500
 DB 1441 AGGACAGATCAAGGCAACAAAGCCTTTAGTGAATTTGCTTGTGTCGCGCAGAG 1500
 QY 1501 GATCTGTTCAATTAAGATGATTAACAAATTTGATACCGATTAATTTTATGCTTTC 1560
 DB 1501 GATCTGTTCAATTAAGATGATTAACAAATTTGATACCGATTAATTTTATGCTTTC 1560
 QY 1561 GTGCTGTGCTTGAATCTTAACGCGCATTCATTAACCTTTAAACGATTCGAAATTC 1620
 DB 1561 GTGCTGTGCTTGAATCTTAACGCGCATTCATTAACCTTTAAACGATTCGAAATTC 1620
 QY 1621 ACAGGGGGCAATGATTTGAAACATTAATCACTAAGCGCTAATGCTACTATTCTG 1680
 DB 1621 ACAGGGGGCAATGATTTGAAACATTAATCACTAAGCGCTAATGCTACTATTCTG 1680
 QY 1681 GGAAGGAAGATTTGTTCTAATGAAATTAATTAATTAATTAATTAATTAATTAAT 1740
 DB 1681 GGAAGGAAGATTTGTTCTAATGAAATTAATTAATTAATTAATTAATTAATTAAT 1740
 QY 1741 AAATGCTCAACAGTGTGTTTGGGAAACAGATTAATAATTAACATGAGGCGATTA 1800
 DB 1741 AAATGCTCAACAGTGTGTTTGGGAAACAGATTAATAATTAACATGAGGCGATTA 1800
 QY 1801 ACCTTATTTAATCAACACAGAAATGCTTCTTGGCTATCTTCAAGGTGGTCAAAAT 1860
 DB 1801 ACCTTATTTAATCAACACAGAAATGCTTCTTGGCTATCTTCAAGGTGGTCAAAAT 1860
 QY 1861 TAAAGGCGATTTACCAACAAAGATTAATTTTTCAGGGTGAACCAACCCG 1920
 DB 1861 TAAAGGCGATTTACCAACAAAGATTAATTTTTCAGGGTGAACCAACCCG 1920
 QY 1921 ACAGCTCAATCATTTAATTAATTAATGCTTGTCAAGAAATGAGATTAACCAAGCGGAA 1980
 DB 1921 ACAGCTCAATCATTTAATTAATTAATGCTTGTCAAGAAATGAGATTAACCAAGCGGAA 1980

QY	1981	TTGGTGGGATCAGATGGATCAACCGGATCATTTAAAGCTGAAACCTTCCAAATTAAG	204
Db	1981	TTGGTGGGATCAGATGGATCAACCGGATCATTTAAAGCTGAAACCTTCCAAATTAAG	2040
QY	2041	GCGAAGTGGGTGGTTCCTCGCAATGTTTCTTCAATTAAGGAAATGGAACAGTCGCA	2100
Db	2041	GCGAAGTGGGTGGTTCCTCGCAATGTTTCTTCAATTAAGGAAATGGAACAGTCGCA	2100
QY	2101	ATTAATGCAATGCGCATTTGGTGTGTCGCAATCAACAAATACATTTGACGCGGT	2160
Db	2101	ATTAATGCAATGCGCATTTGGTGTGTCGCAATCAACAAATACATTTGACGCGGT	2160
QY	2161	CAGATGGAACAGATTAAAGACTGTCAAAAAGTGGATTACCGATACAAAGTATTA	2220
Db	2161	CAGATGGAACAGATTAAAGACTGTCAAAAAGTGGATTACCGATACAAAGTATTA	2220
QY	2221	ATTCTATACCAAAAACCAATCAATGCGCTCTTAATTAACGTAAATGCAAGGCA	2280
Db	2221	ATTCTATACCAAAAACCAATCAATGCGCTCTTAATTAACGTAAATGCAAGGCA	2280
QY	2281	ATGTTAAAGTTTACGAAAATTAAATGGCAATGCACTTTAAACAATCACAGCAATTA	2340
Db	2281	ATGTTAAAGTTTACGAAAATTAAATGGCAATGCACTTTAAACAATCACAGCAATTA	2340
QY	2341	CATTAGCAACAATGCCACCCCAATAGGCAATTCGACTTCCGCAATTCACGCA	2400
Db	2341	CATTAGCAACAATGCCACCCCAATAGGCAATTCGACTTCCGCAATTCACGCA	2400
QY	2401	CGGTGATATATGCAAACTTGAAACGCTATGTGACTTTAGAGATTCAGTCAATTTCT	2460
Db	2401	CGGTGATATATGCAAACTTGAAACGCTATGTGACTTTAGAGATTCAGTCAATTTCT	2460
QY	2461	TAAAAACGCCATTTTTCGACCAAAATTCAGGAGACAAAGGCAACAAGTACGTTGG	2520
Db	2461	TAAAAACGCCATTTTTCGACCAAAATTCAGGAGAGAAAGGCAACAAGTACGTTGG	2520
QY	2521	AAATATCGACTTGGACAAATGCTTACGATCTACTACATTCGAAATTTTAACGCTAATTA	2580
Db	2521	AAATATCGACTTGGACAAATGCTTACGATCTACTACATTCGAAATTTTAACGCTAATTA	2580
QY	2581	GTACGATACGCTTAAATTAATGACTTATACACTTACGCTCAAAATATGCAAGTGGCGGT	2640
Db	2581	GTACGATACGCTTAAATTAATGACTTATACACTTACGCTCAAAATATGCAAGTGGCGGT	2640
QY	2641	CATTAGAGCGAAACCAAGCCCAATCGGAGAAACATCGTTTCAACACAATTGACGTTA	2700
Db	2641	CATTAGAGCGAAACCAAGCCCAATCGGAGAAACATCGTTTCAACACAATTGACGTTA	2700
QY	2701	ATGCTAAATGATGGGCAAGCACTTCCAAATTAATTAATTAATTAATTAATTAATTA	2760
Db	2701	ATGCTAAATGATGGGCAAGCACTTCCAAATTAATTAATTAATTAATTAATTAATTA	2760
QY	2761	GGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2820
Db	2761	GGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2820
QY	2821	CAGGCAAAAGAACCCGAAACCCCTTAGGCAATTAATCTTGGTGAAGCAAAAGATATCAAC	2880
Db	2821	CAGGCAAAAGAACCCGAAACCCCTTAGGCAATTAATCTTGGTGAAGCAAAAGATATCAAC	2880
QY	2881	CGTTATACGATTAAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2940
Db	2881	CGTTATACGATTAAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2940
QY	2941	GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3000
Db	2941	GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3000
QY	3001	AATTGCAACATGATTTAGTAAAGACAGACAGACAGACAAATTAAGAACCCAAACAG	3060
Db	3001	AATTGCAACATGATTTAGTAAAGACAGACAGACAGACAAATTAAGAACCCAAACAG	3060
QY	3061	TTGAACCACTGCTAAACACAAACAGTGTAGCCAAAGTGGCTGCAAGAGACAGCA	3120

Db	3061	TTGAAACCAATGCTAAACACAAACAGGTGACCCAAAGGCGGTCAAGAAAGACGCCAC	3122
Qy	3121	GAACGACGGTTTCGTGATACCTCGCTGATCAAAAGCTGTTAAACCATTAAGAACCCAAAC	3183
Db	3121	GAACGACGGTTTCGTGATACCTCGCTGATCAAAAGCTGTTAAACCATTAAGAACCCAAAC	3180
Qy	3181	AAGCTGAATGATGCTGAAACACAAACAAAGTAAAGCAAAAACAAAAGAGTGGGTGACA	3240
Db	3181	AAGCTGAATGATGCTGAAACACAAACAAAGTAAAGCAAAAACAAAAGAGTGGGTGACA	3240
Qy	3241	AAAGACAGTGTTCCTGATCCCTGCTTGATCAAAAGCCTGTTCCGATTTGAAGACCGCAC	3300
Db	3241	AAAGACAGTGTTCCTGATCCCTGCTTGATCAAAAGCCTGTTCCGATTTGAAGACCGCAC	3300
Qy	3301	TTGAGGTTATGAGACCCCAACAGCATGCGAAAAAGATGCTTACGTCAGAAAGACCGCG	3360
Db	3301	TTGAGGTTATGAGACCCCAACAGCATGCGAAAAAGATGCTTACGTCAGAAAGACCGCG	3360
Qy	3361	AAAAACAACGCAACAAAAAAGCTTGATCAGCCGTTATTCATATGTCGCTTACAGAT	3422
Db	3361	AAAAACAACGCAACAAAAAAGCTTGATCAGCCGTTATTCATATGTCGCTTACAGAT	3422
Qy	3421	TATCTGCAACGTAATACGATGCTTCTGTCAAGATGATTAAGATGCTTTTGTAG	3480
Db	3421	TATCTGCAACGTAATACGATGCTTCTGTCAAGATGATTAAGATGCTTTTGTAG	3480
Qy	3481	ATCAGACCAATGCGCGGTGAGCAAAATATGCAAGATTAAGAGCGTATGATTCG	3540
Db	3481	ATCAGACCAATGCGCGGTGAGCAAAATATGCAAGATTAAGAGCGTATGATTCG	3540
Qy	3541	ATGCGTTCGCTTATCAGCAGCAACAAACGAATTAAGTCAATGAGCGGTGCAAAAG	3600
Db	3541	ATGCGTTCGCTTATCAGCAGCAACAAACGAATTAAGTCAATGAGCGGTGCAAAAG	3600
Qy	3601	CCTTACGTAATGAGCAATGCGGCAAGTTTCTCGCAAGCCGTTACGATTAATACCTTG	3660
Db	3601	CCTTACGTAATGAGCAATGCGGCAAGTTTCTCGCAAGCCGTTACGATTAATACCTTG	3660
Qy	3661	ATGAAACAGGTTAAATAATCAGCGCACTTAACGATGATGTCGGGTTTGCCCAATCAAT	3720
Db	3661	ATGAAACAGGTTAAATAATCAGCGCACTTAACGATGATGTCGGGTTTGCCCAATCAAT	3720
Qy	3721	GGGCGAGATTACATTTGGTGTAAACGTGCGAAGCGGCAATCAGTGCAGTAATGCGTG	3780
Db	3721	GGGCGAGATTACATTTGGTGTAAACGTGCGAAGCGGCAATCAGTGCAGTAATGCGTG	3780
Qy	3781	AAGAAACAAAGCCGAAATAATCATGAAAGCGATTAATTAAGCGTGAATGCAAGTATC	3840
Db	3781	AAGAAACAAAGCCGAAATAATCATGAAAGCGATTAATTAAGCGTGAATGCAAGTATC	3840
Qy	3841	AGTTCGGTTAAGGCAATGCGCAATGAGCTTAATTTGAGAGTTAAATGCTATTTATG	3900
Db	3841	AGTTCGGTTAAGGCAATGCGCAATGAGCTTAATTTGAGAGTTAAATGCTATTTATG	3900
Qy	3901	AACGTGAAATTAATCATCTGAGAGAGTGAAGTGAAGAACCGCTAGCCTGCATTTAATC	3960
Db	3901	AACGTGAAATTAATCATCTGAGAGAGTGAAGTGAAGAACCGCTAGCCTGCATTTAATC	3960
Qy	3961	GCTAATAAGCTGCGCATTTGAGTGAATTAATCACTTACCGACAGTAATATACAGCTTA	4020
Db	3961	GCTAATAAGCTGCGCATTTGAGTGAATTAATCACTTACCGACAGTAATATACAGCTTA	4020
Qy	4021	AACCTTATTTCTCGTCATTTATGTAATTTTCAAGGCTTAAGTCAAAACACGAGTA	4080
Db	4021	AACCTTATTTCTCGTCATTTATGTAATTTTCAAGGCTTAAGTCAAAACACGAGTA	4080
Qy	4081	ATCTCAAGGTGTTGCAACACCAATTTGACGTTATGCGCAAAAAGAGTGGATTAAAG	4140
Db	4081	ATCTCAAGGTGTTGCAACACCAATTTGACGTTATGCGCAAAAAGAGTGGATTAAAG	4140
Qy	4141	CAGAAATTTACATTCGAATTCGCTTTATATCTCAAAATCTCAAGGTTCAACAATCG	4200
Db	4141	CAGAAATTTACATTCGAATTCGCTTTATATCTCAAAATCTCAAGGTTCAACAATCG	4200

Db 4141 CAGAAATTTTACATTTCCAAATTTCCGCTTTATATCTCAAAATCTCAAGTTCAACTCG 4200
 Qy 4201 GCAACAGCAAAATGCGCGCTGAAATATGCGCTATCGTTGGTAAATCAACATATTTT 4260
 Db 4201 GCAACAGCAAAATGCGCGCTGAAATATGCGCTATCGTTGGTAAATCAACATATTTT 4260
 Qy 4261 ATCGTTATTTGATTAACAGAGTGGGTCAATCAATCCACCTTTTATTCATATAT 4319
 Db 4261 ATCGTTATTTGATTAACAGAGTGGGTCAATCAATCCACCTTTTATTCATATAT 4319

RESULT 4

PCT-US03-17092-520/c
 ; Sequence 520, Application PC/TUS0317092
 ; GENERAL INFORMATION:
 ; APPLICANT: BAKLETZ et al.
 ; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF H. INFLUENZAE
 ; FILE REFERENCE: 28335/38815
 ; CURRENT APPLICATION NUMBER: PCT/US03/17092
 ; NUMBER OF SEQ ID NOS: 588
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 520
 ; LENGTH: 8439
 ; TYPE: DNA
 ; ORGANISM: H. influenzae
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (8305)..(8309)
 ; OTHER INFORMATION: n = a, c, g, or t
 ; NAME/KEY: misc_feature
 ; LOCATION: (8427)..(8427)
 ; OTHER INFORMATION: n = a, c, g, or t
 PCT-US03-17092-520

Query Match 73.9%; Score 3193.4; DB 1; Length 8439;
 Best Local Similarity 85.3%; Pred. No. 0;
 Matches 3710; Conservative 0; Mismatches 576; Indels 61; Gaps 11;

Qy 1 TCATAGTGTGTTAACTA-GTATTTTAAAGCAAAATTTACTTAAATTAATTAAGATT 59
 Db 6039 TCATAGTGTGTTAACTAAGTATTTTAAAGCAAAATTTACTTAAATTAATTAAGATT 5980
 Qy 60 ATGAAAAAACTGTATTCGTCTTAATTTTAAACCGCTGCACTTCAATAGGATAGTA 119
 Db 5979 ATGAAAAAACTGTATTCGTCTTAATTTTAAACCGCTGCACTTCAATAGGATAGTA 5920
 Qy 120 TCGCAACCGTGGCTGCTGCACTTATTTGGGATTTGATACCAATATATTCGATTTT 179
 Db 5919 TCGCAACCGTGGCTGCTGCACTTATTTGGGATTTGATACCAATATATTCGATTTT 5860
 Qy 180 GCCGAGATTAAGGAGAGTTCACAGTTGGGCTCAAAATTTAAGTTTATTAACAAACAA 239
 Db 5859 GCTGAGATTAAGGAGAGTTCACAGTTGGGCTCAAAATTTAAGTTTATTAACAAACAA 5800
 Qy 240 GGGCAATAGTTGGCAATCAATGACAAAGCCCGATGATATATTTTCTGTAAGTCA 299
 Db 5799 GGGCAATAGTTGGCAATCAATGACAAAGCCCGATGATATATTTTCTGTAAGTCA 5740
 Qy 300 CGTAAAGCGGTGGCGCTTGTGTAATTAATTAATTTGAGCGTGGCACTAATAGCTA 359
 Db 5739 CGTAAAGCGGTGGCGCTTGTGTAATTAATTAATTTGAGCGTGGCACTAATAGCTA 5680
 Qy 360 GGAATATACAGATGTTGATTTGTGCAAGAGGAAACCCCGATCAACATCGTTTACT 419
 Db 5679 GGAATATACAGATGTTGATTTGTGCAAGAGGAAACCCCGATCAACATCGTTTACT 5620
 Qy 420 TATAAGATTGTAAACGAAATTAACAAAAAAATTAATTAATTCATCTTATAGAGAGAT 479
 Db 5619 TATCAATATGTAAACGAAATTAATTAACAACTGATCAAGAGATCATATATAGGGAT 5560
 Qy 480 TACCAATATCCAGATTACATTAATTCGTTACAGAGGCGCTCAATGATATGACTTCG 559

Db 5559 TACCAATATCCAGATTACATTAATTCGTTACAGAGGCGCTCAATGATATGACTTCG 5500
 Qy 540 AATATGATGCAAGTATTAATCAATGAAACAAATTAATTCAGAAACGTTGCTATCGCC 593
 Db 5499 AGTATGAAAGGACAGACTATATGCAACAAAGAAATTAATTCAGAAACGTTGCTATCGCC 5440
 Qy 600 TCTGACGCGAGTTTGTGCAATATGATCAAGCAAAAGGACCAAGTTGCGGTGATAT 659
 Db 5439 TCTGATGCAATTTGTGCAATATGATCAAGCAAAAGGACCAAGTTGCGGTGATAT 5380
 Qy 660 CATATCTGACAGCTGGCAATACACATATGCGTGGAGCGTATATGATATGATAT 719
 Db 5379 CATATCTGACAGCTGGCAATACACATATGCGTGGAGCGTATATGATATGATAT 5320
 Qy 720 TTGGAGGCGCATGTTGTGTAAGGGGAGAAATGTCATTCGATTCGACGCTCAAG 779
 Db 5319 TTAGTGGCGATGAGTAAAGGAAATGATATGACCTATGCGACGAGATCATTT 5260
 Qy 780 GGGGACAGTGTCTCCGATGTTTATATGATGCTGAAAAAATAATGTTAAAT 839
 Db 5259 GGGGACAGTGTCTCCGATGTTTATATGATGCTGAAAAAATAATGTTAAAT 5200
 Qy 840 GGGATTTAGGGGAAAGCAACCTTTTGAAGCAAAAGAAATGGTTTCAATTGTTGCC 899
 Db 5199 GGTGATATGCGTACTGGAACCCATCTTATGAGAGAAATCTTTCAGCTAGTCCGT 5140
 Qy 900 AATCTTATTTTATGTAATTTTTCGAAAGAGAT-TTACATCATCTTATACCCGAGC 958
 Db 5139 AAAAAATTTTATGTAATTTTCTCTGATTTTAAACACATATTTGATTCAG 5080
 Qy 959 TGGTATGAGCTGACCAATATAGTGA-----ATGATATGTCAGCGGCTATAT 1010
 Db 5079 ACTCCGTAATATCAAAATATGAGCATACCACTTATTCATTAATAATATGATGCTAGCC 5020
 Qy 1011 ACTCAAAATAGAGATACATCAAGAAATTAATTAATTCGTTAGCAATATAGTTACT 1070
 Db 5019 ACTATATATCTGAGCTGTTAAATTTTATTAAGTCAAACTTGTATTAATTAATTC 4960
 Qy 1071 TGGAAAGAGAGTAAAGTCTATATGCTATATGAGAGAGCTTAATTTATTCGCA 1130
 Db 4959 ATGAAAGAGAGAGCTGCTGCTATATCT-----AATGCTGCTTAACATTAATTCGCC 4906
 Qy 1131 CGTTTAAACATGAGAAAGCTATATTTATGATCAAAACAAAGATCATTAATCTTC 1190
 Db 4905 GATTAAGTACGCGGCAAAACCTTATTTATGATTAAGAAAGGCGTCTAATCTTC 4846
 Qy 1191 GATCTGACATTAACAAAGGCGGCTGCTTATTTTGAAGGTAATTTTACAGATTC 1250
 Db 4845 AGTAAATATATGATCAAGGTCAGGCGGTTGATTTTGAAGGAGATTTTATAGTTTCG 4786
 Qy 1251 CCAATCTTACCAAACTTGGCAAGCTGCTCATATGTAAGTGAATAATGACACGTT 1310
 Db 4785 CTTACCACTATGTAAGCTGTAAGAGGCGGCTGATATCTATGATGTAAGACACGTT 4726
 Qy 1311 ACTGGAAGTAAATGCGTGGACATGATGCACTTCTTAATTTGTAAGAGACATTTG 1370
 Db 4725 ATTGGAAGTCAATGCGTGAAGAAATGATGCTTCTTAATTTGTAAGAGACATTTG 4666
 Qy 1371 CAGCTTCAAGCAAAAGGCAAAATTAAGTTGATGAGCTGTAAGATGAT 1430
 Db 4665 CATATTTCAAGCAAAAGGCAAAATTAAGTTGATGAGCTGTAAGATGAT 4606
 Qy 1431 TTGAGAGCAGCAGGCAAGTCAAGGCAAAACAAAGCCTTATGTAAGTGGCTTGGT 1490
 Db 4605 TTGAGACAAAGAGAGCTATCAATTAATAAACAGCCTTATGTAAGTGGT 4546
 Qy 1491 AGGCGAGAGGAGCTTCAATTAACAGATGATTAACAAATTTGATACGATTAATTTAT 1550
 Db 4545 AGGCGAGAGGAGCTTCAATTAACAGATGATTAACAAATTTGATACGATTAATTTAT 4486
 Qy 1551 TTGGCTTGTGCTGCTGCTTATGATTTAGCGGATTTCACTTAACCTTAACGCTATC 1610

Db 4485 TTGGCTTTCGCTGCTGCTTACATCTTAATGAGCAATTAATTAACCTTTAAAGCTATC 4426
 Qy 1611 CAAATAACGACGAGGAGGCAATATGTTGACCATTAATACCACTCAAGCCGTATATGC 1670
 Db 4425 CAAAATACGACGAGGAGGCAATATGTTGACCATTAATACCACTCAAGCCGTATATGC 4366
 Qy 1671 ACTATTAATGAGGACGAAAGCAATGTTCTACCTTAATGAAATTAATTAATTAACCTAT 1730
 Db 4365 ACTATTAATGAGGACGAAAGCAATGTTCTACCTTAATGAAATTAATTAATTAACCTAT 4306
 Qy 1731 TACGAAAGAAATTTGCTTACCAAGCTTGTGTTGGGAAACAGATTAATTAATTAACCAAT 1790
 Db 4305 TACGAAAGAAATTTGCTTACCAAGCTTGTGTTGGGAAACAGATTAATTAATTAACCAAT 4246
 Qy 1791 GAGGCAATTAACCTTATTAATTAACCAAGCAAGAAAGCTTGTGTTGGGAAACCAAT 1850
 Db 4245 GAGGCAATTAACCTTATTAATTAACCAAGCAAGAAAGCTTGTGTTGGGAAACCAAT 4186
 Qy 1851 GGTACAAATTAATAAGGCGATTAATCCCAACCAAAAGTAACTATTTTCAAGCGTACA 1910
 Db 4185 GGTACAAATTAATAAGGCGATTAATCCCAACCAAAAGTAACTATTTTCAAGCGTACA 4126
 Qy 1911 CCGACACCGACGCTTACCAATCATTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1970
 Db 4125 CCGACACCGACGCTTACCAATCATTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4072
 Qy 1971 CAAAGCAAAATTTGTGAGGATCAAGATTTGATCAAGCTTACCTTAATTAATTAATTAAT 2030
 Db 4071 CAAAGCAAAATTTGTGAGGATCAAGATTTGATCAAGCTTACCTTAATTAATTAATTAAT 4012
 Qy 2031 CAAATTAAGGCGAAAGTGTGCTGCTTCTGCAATGTTTCTCAATTAAGGAAATTTGG 2090
 Db 4011 CAAATTAAGGCGAAAGTGTGCTGCTTCTGCAATGTTTCTCAATTAAGGAAATTTGG 3952
 Qy 2091 AAGCTCAGCAATTAAGCAATGCAATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2150
 Db 3951 AAGCTCAGCAATTAAGCAATGCAATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3892
 Qy 2151 TGCAGCGCTTCAAGATTTGCAAGATTTGCAAGATTTGCAAGATTTGCAAGATTTGCAAGAT 2210
 Db 3891 TGCAGCGCTTCAAGATTTGCAAGATTTGCAAGATTTGCAAGATTTGCAAGATTTGCAAGAT 3832
 Qy 2211 AAGCTTATTAATTTCAAGCAAAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2270
 Db 3831 AAGCTTATTAATTTCAAGCAAAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3772
 Qy 2271 GCAAGCGCAATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 2330
 Db 3771 GCAAGCGCAATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 3712
 Qy 2331 AGCCAAATTAATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2390
 Db 3711 AGCCAAATTAATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 3652
 Qy 2391 TCAATTCGCAAGCTGATTAATGCAAACTTGAACGTTAATGCAATTAATTAATTAATTAAT 2450
 Db 3651 TCAATTCGCAAGCTGATTAATGCAAACTTGAACGTTAATGCAATTAATTAATTAATTAAT 3592
 Qy 2451 CAATTTCTTTAAATAAAGCCATTTTTCGCAAAATTTAAGGAGAAAGGAGGAGGAGGAGGAGG 2510
 Db 3591 CAATTTCTTTAAATAAAGCCATTTTTCGCAAAATTTAAGGAGAAAGGAGGAGGAGGAGGAGG 3532
 Qy 2511 GTGACGTTGAAATGCACTTGAACATGCTTGAACATGCTTGAACATGCTTGAACATGCTTGAAC 2570
 Db 3511 GTGACGTTGAAATGCACTTGAACATGCTTGAACATGCTTGAACATGCTTGAACATGCTTGAAC 3472
 Qy 2571 CTAAATTAAGTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2630
 Db 3471 CTAAATTAAGTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3412
 Qy 2631 -----CGTCCGCTTCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2684
 Db 3411 CGCGCTGCTGCTGCTTCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3352
 Qy 2685 AACACATGACAGTAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2744
 Db 3351 AACACATGACAGTAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3292
 Qy 2745 TTAATTTGCTTAAATAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2804
 Db 3291 TTAATTTGCTTAAATAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3232
 Qy 2805 TTAATTTGCTTAAATAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2864
 Db 3231 TTAATTTGCTTAAATAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3172
 Qy 2865 AGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2924
 Db 3171 AGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3112
 Qy 2925 GATCAGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2984
 Db 3111 GATCAGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3052
 Qy 2985 CCAATTAATAAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3044
 Db 3051 CCAATTAATAAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2992
 Qy 3045 TTAAGAGGCAAAAGGTTGAACGCTGCTTAAACCAAAAGGAGGAGGAGGAGGAGGAGGAGG 3104
 Db 2991 TTAAGAGGCAAAAGGTTGAACGCTGCTTAAACCAAAAGGAGGAGGAGGAGGAGGAGGAGG 2932
 Qy 3105 TTAAGAGGCAAAAGGTTGAACGCTGCTTAAACCAAAAGGAGGAGGAGGAGGAGGAGGAGG 3164
 Db 2931 TTAAGAGGCAAAAGGTTGAACGCTGCTTAAACCAAAAGGAGGAGGAGGAGGAGGAGGAGG 2881
 Qy 3165 GCATTAAGAGGCAAAAGGTTGAACGCTGCTTAAACCAAAAGGAGGAGGAGGAGGAGGAGG 3224
 Db 2881 GCATTAAGAGGCAAAAGGTTGAACGCTGCTTAAACCAAAAGGAGGAGGAGGAGGAGGAGG 2821
 Qy 3225 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3275
 Db 2821 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2761
 Qy 3275 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3332
 Db 2761 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2701
 Qy 3332 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3392
 Db 2701 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2650
 Qy 3392 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3452
 Db 2650 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2590
 Qy 3452 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3512
 Db 2590 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2530
 Qy 3512 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3572
 Db 2530 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2473
 Qy 3572 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3632
 Db 2473 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2413
 Qy 3632 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3692
 Db 2413 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2353
 Qy 3692 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3752
 Db 2353 AAAAAAGGCGCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2293

Qy 3753 ACGGATACGTCGATGAAATGCGTGAAGAAAGCCGAAAAATTCATCGAAAGCG 3812
 Db 2292 GCGGAAATAGTGGGATGAAATGCGTGAAGAAAGCCGAAAAATTCATCGAAAGCG 2233
 Qy 3813 AATAATATGCGTGAATGCAATGATACATTCCTGTTTGGGCAATTTGAGCTT 3872
 Db 2232 AATAATATGCGTGAATGCAATGATACATTCCTGTTTGGGCAATTTGAGCTT 2173
 Qy 3873 TATTTGGAGTAAATGCTATTTTATGAAAGTAAATTTAAATTCATCGAGAGTGA 3932
 Db 2172 TATTTGGAGTAAATGCTATTTTATGAAAGTAAATTTAAATTCATCGAGAGTGA 2113
 Qy 3933 GTGAAAAGCTTACCTGCTGATTTATGCTATTAATGCTGCAATTCGATTAACA 3992
 Db 2112 GTGAAAAGCTTACCTGCTGATTTATGCTATTAATGCTGCAATTCGATTAACA 2053
 Qy 3993 TTTACTCGACATTAATTAATGAGGTTAAACCTTATTTCTGCAATTTATGTAATTT 4052
 Db 2052 TTTACTCGACATTAATTAATGAGGTTAAACCTTATTTCTGCAATTTATGTAATTT 1993
 Qy 4053 TCAAAAGCTTACCTGCTGATTTATGCTATTAATGCTGCAATTCGATTAACA 4112
 Db 1992 TCAAAAGCTTACCTGCTGATTTATGCTATTAATGCTGCAATTCGATTAACA 1933
 Qy 4113 TATTCGCAAAAAGAGTGGGATTAAGGCAAAATTTTCAATTTCCGCTTTT 4172
 Db 1932 TATTCGCAAAAAGAGTGGGATTAAGGCAAAATTTTCAATTTCCGCTTTT 1873
 Qy 4173 ATCTCAAAATCTCAAGTTTCACATCTCGGCAAAAGCAAAATGCGCGCTGAATGCGC 4232
 Db 1872 ATCTCAAAATCTCAAGTTTCACATCTCGGCAAAAGCAAAATGCGCGCTGAATGCGC 1813
 Qy 4233 TATGCTGTAATAATCAATATTTTATGCTTATGTAATAAAGGCTGAGATC 4292
 Db 1812 TATGCTGTAATAATCAATATTTTATGCTTATGTAATAAAGGCTGAGATC 1153
 Qy 4293 AGATCCACCTTTTATTTCCATTAAT 4319
 Db 1752 AATCCACCTTTTATTTTCCATTAAT 1726

RESULT 5
 US-60-453-134-520/c
 ? Sequence 520, Application US/60453134
 ? GENERAL INFORMATION:
 ? APPLICANT: BAXALETZ et al.
 ? TITLE OF INVENTION: GENES OF AN ORTIS MEDIA ISOLATE OF H. INFLUENZAE
 ? FILE REFERENCE: 2835/38815
 ? CURRENT APPLICATION NUMBER: US/60/453,134
 ? NUMBER OF SEQ ID NOS: 588
 ? SOFTWARE: PatentIn version 3.2
 ? SEQ ID NO 520
 ? LENGTH: 8439
 ? TYPE: DNA
 ? ORGANISM: H. influenzae
 ? FEATURE:
 ? NAME/KEY: misc feature
 ? LOCATION: (8309)..(8309)
 ? OTHER INFORMATION: n = a, c, g, or t
 ? FEATURE:
 ? NAME/KEY: misc feature
 ? LOCATION: (8427)..(8427)
 ? OTHER INFORMATION: n = a, c, g, or t
 US-60-453-134-520

Query Match 73.9%; Score 3193.4; DB 102; Length 8439;
 Best Local Similarity 85.3%; Pred. No. 0;
 Matches 710; Conservative 0; Mismatches 576; Indels 61; Gaps 11;
 Qy 1 TCAATAGTGTAACTA-GTATTTTAAATGCAAAATTAATTAATTAATTAATTAAT 59
 Db 6039 TCAATAGTGTAACTAAGTATTTTAAATGCAAAATTAATTAATTAATTAATTAAT 5980

Qy 60 AAAAAAAAACTGATTTTCGCTTAATTTTAAACCGCTTGCAATTCATAGGATAGTA 115
 Db 5979 AAAAAAAAACTGATTTTCGCTTAATTTTAAACCGCTTGCAATTCATAGGATAGTA 5920
 Qy 120 TCGAAGCGTGGGCTGCTCACTTATTTGGGATGATTAATTAATTAATTAATTAAT 179
 Db 5919 TCGAAGCGTGGGCTGCTCACTTATTTGGGATGATTAATTAATTAATTAATTAAT 5860
 Qy 180 GCGGAAATTAAGGAAAGTTCAAGTTTGGGCTCAAAATTTAAAGTTATTAACAA 239
 Db 5859 GCGGAAATTAAGGAAAGTTTCAAGTTTGGGCTCAAAATTTAAAGTTATTAACAA 5800
 Qy 240 GGGCAATTTGTTGCAATCAATGAATAAACCCCGATGATTTTTCGTGTGTCA 299
 Db 5799 GGGCAATTTGTTGCAATCAATGAATAAACCCCGATGATTTTTCGTGTGTCA 5740
 Qy 300 GGTAAAGCGGTGAGCTTGTGTTAAATCAATATTTTGTAGCGTGGCAATACATA 359
 Db 5739 GGTAAAGCGGTGAGCTTGTGTTAAATCAATATTTTGTAGCGTGGCAATACATA 5680
 Qy 360 GATATACGATGTTGATTTTGTGCAAGGAAACAAACCCGATCAATCGTTTACT 419
 Db 3679 GATATACGATGTTGATTTTGTGCAAGGAAACAAATCCGATCAATCGTTTACT 5620
 Qy 420 TATAGATTTAAAGCAATTAATAAATAAATAATTTACATCTTAATAGAGAGAT 479
 Db 5619 TATCAATTTAAAGCAATTAATAAATAAATAATTTACATCTTAATAGAGAGAT 5560
 Qy 480 TACATTAATCAAGATTAATTAATTTGTTACAGAAAGCTTCAATTTATAGACTTG 539
 Db 5559 TACATTAATCAAGATTAATTAATTTGTTACAGAAAGCTTCAATTTATAGACTTG 5500
 Qy 540 AATATGATGCAATTTATTTTCAATGAATGAACAAATATTCAGAAAGTTGTATCGC 599
 Db 5499 AATATGATGCAATTTATTTTCAATGAATGAACAAATATTCAGAAAGTTGTATCGC 5440
 Qy 600 TCGACGCGAGTTTGGGAAATGATGACAGCAAGGCAAGGCTGCGGTGAT 659
 Db 5439 TCGACGCGAGTTTGGGAAATGATGACAGCAAGGCAAGGCTGCGGTGAT 5380
 Qy 660 CATTAATGACGCTGCAATACACAAATAGCGTGGAGCAATTAATTTGAT 719
 Db 5379 CATTAATGACGCTGCAATACACAAATAGCGTGGAGCAATTAATTTGAT 5320
 Qy 720 TTGGAGGCGATGTTGTAAGCGGAGATTTGTCAATTAACGATTTGAGCTGAAG 779
 Db 5319 TTGGAGGCGATGTTGTAAGCGGAGATTTGTCAATTAACGATTTGAGCTGAAG 5260
 Qy 780 GGGGACAGTGTCTCGAGTTTATTTATGATGCTGAAACAAAAATGTTAATTAAT 839
 Db 5259 GGGGACAGTGTCTCGAGTTTATTTATGATGCTGAAACAAAAATGTTAATTAAT 5200
 Qy 840 GGGATATTAAGGAAAGCAACCTTTTGAAGCAAAAGAAATGCTTCAATTTGCTGC 899
 Db 5199 GGGATATTAAGGAAAGCAACCTTTGTAAGCAAAAGAAATGCTTCAATTTGCTGC 5140
 Qy 930 AATCTTATTTTGAATTAATTTTGAAGAGAT-TTACATACATCACTTACACCGGAC 958
 Db 5139 AATCTTATTTTGAATTAATTTTGAAGAGAT-TTACATACATCACTTACACCGGAC 5080
 Qy 959 TGGTATGAGTGTACCAATTAAGGAA-----ATGATTAATGCTGAGGCTGATA 1010
 Db 5079 TGGTATGAGTGTACCAATTAAGGAA-----ATGATTAATGCTGAGGCTGATA 5020
 Qy 1011 ACTCGAATTAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1370
 Db 5019 ACTCGAATTAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4960
 Qy 1071 TTGAAGAGAGATTAAGTCTTAATCTGATTAATTAATTAATTAATTAATTAATTTTCA 1130
 Db 4959 TTGAAGAGAGATTAAGTCTTAATCTGATTAATTAATTAATTAATTAATTAATTTTCA 4906

QY 1131 CGTTAAACATGAGAAAGCTATATTTATGATCAAAACAGATCATATATCTC 1190
 Db 4905 GCATTTAGTCAAGCGAAACCTTTATTTATGATTAAGAAAGGGGTGTAATCTT 4846
 QY 1191 GCATCTGACATTAACCAAGGGGGGGTGTCTTTATTTGAGGTAATTTTAACATCT 1253
 Db 4845 AGTAAATATATCATATCAAGGTGAGGGGGTGTATTTTGAAGAGATTTTAATGTTG 4786
 QY 1251 CCAATTTCTAACAACTTGCAAGAGCTGGATATAGTAAGTAAATAGACCGTT 1310
 Db 4785 CTAACCATATGAAAGTGAAAGGGGGGGGATATATCATGATGAAAGACCGTA 4726
 QY 1311 ACTTGGAAGTAATGGCTGGACATGATGCACTTTCTAAATTTGGTAAGGACATG 1370
 Db 4725 ATTTGGAAGTCAATGGCTGAGAAATGATGCGCTTTCAAAATTTGGTAAGGACATG 4666
 QY 1371 CAGCTTCAAGCAAGAGGGAATTAAGGTGATGATGATGATGATGATGATGATGATG 1430
 Db 4665 CATATTCAGCAAAAGGTAAATCTTAGGCTCATTTAGGCTGATGATGATGATGATG 4606
 QY 1431 TTGAGAGCAGAGCAGACATCAAGGCAAAACAGCTTTAGTAAATGGCTTGGT 1490
 Db 4605 TTAGAACACAGAGAGAGATCAATTAATAACAGCTTTAGTAAATGGATTTGGT 4546
 QY 1491 AGCGGAGAGGAGCTGTCTATTAAGATGATTAACATTTGATACGATTAATTTAT 1550
 Db 4545 AGCGGAGAGGAGCTGTCTATTAAGATGATTAACATTTGATACGATTAATTTAT 4486
 QY 1551 TTGGCTTGT 1610
 Db 4485 TTGGCTTGT 4426
 QY 1611 CAAATTAAGCAAGAGGAGGCAATGATTTGAAACATTAATCAATCAATCAATCAAT 1670
 Db 4425 CAAATTAAGCAAGAGGAGGCAATGATTTGAAACATTAATCAATCAATCAATCAAT 4366
 QY 1671 ACTTATCTGGAGAGAGAGATTTGTCTTCTACTTATGAAATATTTATTAATCTGAT 1730
 Db 4365 ACTTATCTGGAGAGAGAGATTTGTCTTCTACTTATGAAATATTTATTAATCTGAT 4306
 QY 1731 TACGAAAGAAATTTGCTCAACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1790
 Db 4305 TACGAAAGAAATTTGCTCAACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4246
 QY 1791 GGGGATTAACCTTATTTATTAACCAACACAGAGATGATCTTTGCTTCAAGT 1850
 Db 4245 GAGGATTAACCTTATTTATTAACCAACACAGAGATGATCTTTGCTTCAAGT 4186
 QY 1851 GGTACAAATTTAAAGGAGATTAACCAACAGAGATGATCTTTGCTTCAAGT 1910
 Db 4185 GGTACAAATTTAAAGGAGATTAACCAACAGAGATGATCTTTGCTTCAAGT 4126
 QY 1911 CCGACACGACGCTCAATCATTTAAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1970
 Db 4125 CCGACACGACGCTCAATCATTTAAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4072
 QY 1971 CAAAGGAAATTTGT 2030
 Db 4071 CAAAGGAGATGATTAAGT 4012
 QY 2031 CAAATTAAGGAGAGATGAGT 2090
 Db 4011 CAAATTAAGGAGAGATGAGT 3952
 QY 2091 ACAGTCAAGATTAAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2150
 Db 3951 ACATCAAGATTAAGCAAGT 3892
 QY 2151 TGAAGCGCTTCAATGAGATGAGATTAAGCAATTTGATGATGATGATGATGATGATG 2210
 Db 3891 TGAAGCGCTTCAATGAGATGAGATTAAGCAATTTGATGATGATGATGATGATGATG 3832
 QY 2211 AAGTATTAATTTCTATAACAAACAGCAATTAAGTGTGTGTGTGTGTGTGTGTGTGT 2270

Db 3881 AAGTATTAATTTCTATAACCAACCAATTAATGATGATGATGATGATGATGATGATG 3772
 QY 2271 GCAAGCGCAATGTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 2330
 Db 3771 GCAAGCGCAATGTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 3712
 QY 2331 AGCGATTAATTTAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2390
 Db 3711 AGCGATTAATTTAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3652
 QY 2391 TCAATGCAAGCGT 2450
 Db 3651 TCAATGCAAGCGT 3592
 QY 2451 CAATTTCTTTAAACCAAGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2510
 Db 3591 CAATTTCTTTAAACCAAGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3532
 QY 2511 GTGAGCTGT 2570
 Db 3531 GTGAGCTGT 3472
 QY 2571 CTAAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2630
 Db 3471 CTAAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3412
 QY 2631 -----GCTGCGCTGT 2684
 Db 3411 GCGGCTGT 3352
 QY 2685 AACATTAAGCAATTAAGT 2744
 Db 3351 AACATTAAGCAATTAAGT 3292
 QY 2745 TATTTGCTTAAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2804
 Db 3291 TATTTGCTTAAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3232
 QY 2805 TATCTGT 2864
 Db 3231 TATCTGT 3172
 QY 2865 AGCAATTAATTAAGT 2924
 Db 3171 AGCAATTAATTAAGT 3112
 QY 2925 GATGAGGT 2984
 Db 3111 GATGAGGT 3052
 QY 2985 CCAATTAAGGAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAAT 3044
 Db 3051 CCAATTAAGGAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAAT 2992
 QY 3045 TTAGAGGCAAAAGATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAAT 3104
 Db 2991 TTAGAGGCAAAAGATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAAT 2881
 QY 3105 TTAAG 2932
 Db 2881 TTAAG 2821
 QY 3165 GATTTAG 3224
 Db 2821 GATTTAG 2761
 QY 3225 AAAAAAGT 3275
 Db 2761 AAAAAAGT 2711
 QY 3275 AAAAAAGT 3332
 Db 2711 AAAAAAGT 3275

```

? Sequence 8, Application PC/TUS0305226
? GENERAL INFORMATION:
? APPLICANT: St. Gene, Joseph W.
? TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
? FILE REFERENCE: A-59941-1/RET/DCF/DHR
? CURRENT APPLICATION NUMBER: PCT/US03/05226
? CURRENT FILING DATE: 2003-02-18
? PRIOR APPLICATION NUMBER: US 08/236,791
? PRIOR FILING DATE: 1994-10-25
? PRIOR APPLICATION NUMBER: US 09/839,996
? PRIOR FILING DATE: 2001-04-20
? NUMBER OF SEQ ID NOS: 58
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 8
? LENGTH: 4305
? TYPE: DNA
? ORGANISM: Haemophilus influenzae
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1702)..(1702)
? OTHER INFORMATION: "n" at position 1702 can be any base.
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(4305)
? OTHER INFORMATION:
? PCT-US03-05226-8

```

Query Match	73.64%	Score 3180.8	DB 1	Length 4305
Best Local Similarity	85.74%	Pred. No. 0:		
Matches 3704	Conservative	0	Mismatches 463	Indels 156
Gaps				9
Qy	60	ATGAAAAAACTGATATCCGCTTAATTTTACCGCTSCATTTCATTAGGATGTA	119	
Db	1	ATGAAAAAACTGATATTCGCTTAATTTTACCGCTSCATTTCATTAGGATGTA	60	
Qy	120	TCGAAAGCTGCGCTCGTCACTATTATTTGGCATTTGATACCAATATATGATTTT	179	
Db	61	TCGAAAGCTGCGCGGCTCACTATTATTTGGCATTTGATACCAATATATGATTTT	120	
Qy	180	GCCGAGATPAAAGGAAATTCACATTTGGGGCTCAAAATATAGGTTATPACAA	239	
Db	121	GCCGAGATPAAAGGAAATTCACATTTGGGGCTCAAAATATAGGTTATPACAA	180	
Qy	240	GGGCAATTAGTTGGCATCATCAATGCAAAAGCCCGCATGATTAATTTTCGTGTA	299	
Db	181	GGGCAATTAGTTGGCATCATCAATGCAAAAGCCCGCATGATTAATTTTCGTGTA	240	
Qy	300	CTPAAAGCGCTGGAGGCTTGGTGAATCAATATTTGGAGCGTGGCATATGTA	359	
Db	241	AGAAATGAGATGCTGCTTACATGCGCATGATATTTGGAGTGGCATATGTA	300	
Qy	360	GGATATACAGATGTTGATTTTGGTSCAGAGGGAAACACCCGATCAATCGTTTACT	419	
Db	301	GGCTATACCATGATGCAATTTGGTCTGAGAGCAAAATCTGATCAATCGTTTACT	360	
Qy	420	TATPAGATTTGAACGAATTAATCTACAAAAGATATTAATCATCTTATGAGACAT	479	
Db	361	TATPAAATTTGAAACGAAATATTAATTAATCAAGTGGAGACCGGTATCTAGTAC	420	
Qy	480	TACCAATATCCAGATTAATTAATTTGTTACAGAGCGCTTCAATGATATGACTTG	539	
Db	421	TACCAATATCCAGATTAATTAATTTGTTACAGATGCGGACCAATGATATGACTTCA	480	
Qy	540	AATATGATGCGATCTTATTCAGATGAAACAAATATTCGAACGTTGCTATACGC	599	
Db	481	CATATGATGCGATTAATGATGCAATTAAGGAAATATCTGAACGATGCGTGGTA	540	
Qy	600	TTCTGAGCGAGTTTGGCGAAATGATCAAGACAAAGGCGACCAAGTGGCGGTGCATAT	659	
Db	541	TTCTGAGATGCTATTTGGGATGACATCAAAACAAAGAACTTATTTATCTGAGGATAT	600	
Qy	660	CATATTCGACACTCGGCAATACACCAATCAAGCGTGGACGGTATGATATTCGAT	719	

```

Db      601 AATTATTACAGCTGGGAATACATATATATAAGCGTAAAGCTGATGATATTCAT 660
Qy      720 TTGGAGGCGATGTCGTATAAGCGGAGATATGTCATTAACCATGAGGCTCAAG 779
Db      661 GTGAGAGGTGATTTGGCAAGTGTGGCATTAATGTCATTAACCATGAGGCTCA 720
Qy      780 GGGGACATGCTCTCCGATGTTTATTATGATGCTGAAAACAAAAATGTTATAT 839
Db      721 GGGGACATGATCTCCATGTTTATTATGATGCTGAAAACAAAAATGTTATAT 780
Qy      840 GGGATATTACGGGAGGCAACCTTTGAGGCAAAAATGCTTTGATTTGCTGTC 899
Db      781 GAGATATGCGGAGGCGGCACTTATACGGGCAATTCGATGATTTCAATTAAC 840
Qy      900 AATCTTATTTATGATAAATTTTGGAAAGATTTACATACATCACTTACCCGAG 959
Db      841 AATCTTCTCTGATGATAATTTATGCAAAATGATCAACAAATGTTTATACCT 900
Qy      960 GGTATATGATGATACAAATTATGATGAAATGATATGTCAGGGCTATATCTAG 1019
Db      901 GGGATGCGCTTTATACATTTCTTAAAGGAGATGAGATGAGATGTTACTTGA 960
Qy      1020 TCGGATACCATCAAAATTAATTAAGCTTATGAGTATGATCTTGAAG 1079
Db      961 ATGGAATACC-----TCGTAAATCTTTAGCGAACAATTTAAATGAGAT 1014
Qy      1080 AAGATTAAGTTATATCTTATGATGATGAGGACTTATTTATCTCCAGTTAA 1139
Db      1015 AAGATATCTGCTATATATACATATATATGCTTATTTATCTCTCAATTAAC 1074
Qy      1140 AATGAGAAATGCTATATTTATGATGATGATGATGATGATGATGATGATGAT 1199
Db      1075 AATGCAAGATATTTATTTTATGATGATGATGATGATGATGATGATGATGAT 1134
Qy      1200 ATTAACCAAGGCGGCTGCTTATTTTGGGATTTTACATGATCTCAATTTCT 1259
Db      1135 ATGATCAAGTGTGAGGCGGCTTATTTTGGGAGGATTTTATGATGCTCA 1194
Qy      1260 AACCATACTTGGCAAGAGCTGATATGATGATGATGATGATGATGATGATGAT 1319
Db      1195 AATGAAATGATGAAAGGCGGCTTATGATGATGATGATGATGATGATGATGAT 1254
Qy      1320 GTAAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379
Db      1255 GTAAAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
Qy      1380 GCGAAAGCGGAAATTAAGTTCATGATGATGATGATGATGATGATGATGATGAT 1439
Db      1315 GCGAAAGCGGAAATTAAGTTCATGATGATGATGATGATGATGATGATGATGAT 1374
Qy      1440 CAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
Db      1375 CAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
Qy      1500 GGGATGCTTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1559
Db      1435 GGGATGCTTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1494
Qy      1560 CGTGTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
Db      1495 CGTGTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1554
Qy      1620 GAGGAGGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
Db      1555 GAGGAGGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1614
Qy      1680 GGGAGCAAAAGCAT---TGTTCATCTATGATGATGATGATGATGATGATGATGAT 1735
Db      1615 GGGAGCAAAAGCAT---TGTTCATCTATGATGATGATGATGATGATGATGATGAT 1674
Qy      1737 AAGAAATGCTTCAACAGGTGCTTGGCAAAAGATTAATTAATTAATTAATTAAT 1796
Db      1675 AAGAAATGCTTCAACAGGTGCTTGGCAAAAGATTAATTAATTAATTAATTAAT 1734

```

```

Qy      1797 TTAACTTATTTAAACCAACCAAGATGATGATGATGATGATGATGATGATGAT 1856
Db      1735 TTAACTTATTTAAACCAACCAAGATGATGATGATGATGATGATGATGATGAT 1794
Qy      1857 AATTTAAAGGCAATTAACCAACCAAGATGATGATGATGATGATGATGATGATGAT 1916
Db      1795 AACTTTAAAGGCAATTAACCAACCAAGATGATGATGATGATGATGATGATGATGAT 1854
Qy      1917 CCGACGCTTACATCAATTTAAATTAACCTTGTGATGATGATGATGATGATGATGAT 1976
Db      1855 CCGACGCTTACATCAATTTAAATTAACCTTGTGATGATGATGATGATGATGATGAT 1914
Qy      1977 GAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2036
Db      1915 GAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1974
Qy      2037 AAGGCGAAGTGGGCTGCTTGTGCAATGCTTCAATGAGGAAATGATGATGATGAT 2096
Db      1975 AAGGCGAAGTGGGCTGCTTGTGCAATGCTTCAATGAGGAAATGATGATGATGAT 2034
Qy      2097 AGCATATGCAAAATGCGCATTTGCTGCTGATGATGATGATGATGATGATGATGAT 2156
Db      2035 AGCATATGCAAAATGCGCATTTGCTGCTGATGATGATGATGATGATGATGATGAT 2094
Qy      2157 CGTTCAGTGGGCAAGATTAACGATGATGATGATGATGATGATGATGATGATGAT 2216
Db      2095 CGTTCAGTGGGCAAGATTAACGATGATGATGATGATGATGATGATGATGATGAT 2154
Qy      2217 ATTAATCTATACCAAAACCAATCAATGATGATGATGATGATGATGATGATGATGAT 2276
Db      2155 ATTAATCTATACCAAAACCAATCAATGATGATGATGATGATGATGATGATGATGAT 2214
Qy      2277 GCGATGTTAAAGTTAGCAAACTTAATGATGATGATGATGATGATGATGATGATGAT 2336
Db      2215 GCGATGTTAAAGTTAGCAAACTTAATGATGATGATGATGATGATGATGATGATGAT 2274
Qy      2337 TTAACTTAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2396
Db      2275 TTAACTTAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2334
Qy      2397 -----GCA 2396
Db      2335 GCAAGGTAATTAATGCACTTATGCGGATGATGATGATGATGATGATGATGATGAT 2394
Qy      2397 -----GCA 2399
Db      2395 ACATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2454
Qy      2400 ACATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2459
Db      2455 ACATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2514
Qy      2460 TTAAATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2519
Db      2515 TTAAATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2514
Qy      2520 GAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2579
Db      2575 GAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2534
Qy      2580 AGTACGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2636
Db      2635 AGTACGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2634
Qy      2637 CGTTCATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2696
Db      2695 CGTTCATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2654
Qy      2697 GTAAATGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2756
Db      2755 GTAAATGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2714

```

Db	3883	GGGAGAAATGCAAGTTATGCTTCCGTTTAGGGCAATGGGATTCAGACCTTATTGGGCT	3942
Qy	3882	GTTAATCGCATTTTATTATGACGTGAAATTTATCAATCGAAGAGTGAAGTGAAACG	3941
Db	3943	GTTAATCGCATTTTATTATGAGTGAATTAATTCATCTGAGGAAGTGAAGTGCAACA	4002
Qy	3942	CTAGCGCTTCATTTAATGCTATATATGCTGGCATTTGGATTTATACATTTACTCG	4001
Db	4003	CCGAGCGTTGCATTTAATGCTATATATGCTGGCATTCGAGTTGATTTATACATTTACCCG	4062
Qy	4002	ACAGATATATCGCGCTTAGCGCTTATTTCTGTGATTTATGTTGATTTTAAACGCT	4061
Db	4063	ACAGATATATCGCGCGTTAGCGCTTATTTCTGTGATTTATGTTGATTTTAAACGCT	4122
Qy	4062	AAAGTAGCAACACCGGTAATCTCACGCGTGGCAACACGATTTGGACGTATTGGCA	4121
Db	4123	AAAGTAGCAACACCTGTAAATATGACATGTGTCAACCAATCTATTGGCGCTTATTGGCA	4182
Qy	4122	AAAGTAGTGGGATTTAAGGACAGAAATTTTACATTTCCAAATTCGCGCTTTATCTCAAA	4181
Db	4183	AAAGTAGTGGGATTTAAGGACAGAAATTTTACATTTTCCAACTTTCCGCTTTATCTCAAAA	4242
Qy	4182	TCTCAAGGTTCCAACTGGGCAACACAGCAAAATGTGGGCGTGAATTTGGGCTATGGTGG	4241
Db	4243	TCTCAAGGTTCCAACTGGTGAACAGCAAAATGTGGGCGTGAATTTGGGCGTATGGTGG	4302
Qy	4242	TAA 4244	
Db	4303	TAA 4305	

```

RESULT 7
US-10-687-046-8
; Sequence 8, Application US/10687046
; GENERAL INFORMATION:
; APPLICANT: St. Geme, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/KRT/DCP/DDR
; CURRENT APPLICATION NUMBER: US/10/687.046
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US/10/080,505
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4305
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1702)..(1702)
; OTHER INFORMATION: "n" at position 1702 can be any base.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4305)
; OTHER INFORMATION:
US-10-687-046-8

```

Query Match	73.6%;	Score 1180.8;	DB 53;	Length 4305;
Best Local Similarity	85.7%;	Pred. No. 0;		
Matches 3704;	Conservative	0;	Mismatches 463;	Indels 156;
			Gaps	9;

Qy 60 AAAAAAAAAAGTATTTCGCTTAATTTTAAACCGGTGCATTTCATAGGAGATGTA 119
Db 1 AAAAAAAAAATGTAATTTCGCTTAATTTTAAACCGGTGCATTTCATAGGAGATGTA 60
Qy 120 TCGCAAGGTGCGGTGCTACACTAATTTTGGGAATGATTACCAATATTATTCGATTTT 179

Db 61 TCAGCAAGCTGGAGGCTCACTTATTGGGATGACCTACCAATATATCTGATTTT 120
Qy 180 GCCGAGATTAAGGAGAGTTCACAGTTG999CTCAAAATTTAGTTTAAACAACA 239
Db 121 GCCGAGATTAAGGAGAGTTCAGTTG999CTCAAAATTTAGTTTAAACAACA 180
Qy 240 GGGCAATTAAGTGGACATCAATAGCAAAAGCCCGCATGATTTCTGTAGTCA 299
Db 181 GGGCAATTAAGTGGACATCAATAGCAAAAGCCCGCATGATTTCTGTAGTCA 240
Qy 300 CGTAAAGCGCTGGAGGCTTGGTGAATATCAATATATTTAGAGGAGCACTAACTA 355
Db 241 ACAAATGAGAGTTGCTCTTACAGGCACTACATATTTAGAGGAGCACTAACTA 300
Qy 360 GATATACAGATTTGATTTTGTGCGAGAGGAAACACCCCGATCAACATCGTTTACT 419
Db 301 GAGTATACAGATTTGATTTTGTGCGAGAGGAAACACCCCGATCAACATCGTTTACT 360
Qy 420 TATTAAGTTGTAAGAAAGAAATTAACATAAAGAAATTAATTTACATCTTTAGAGAGAT 479
Db 361 TATTAAGTTGTAAGAAAGAAATTAATTAATCAAGATGCGAAGACCGCTATAGATGAC 420
Qy 480 TACCATATCAACGATTAATTAATTTGTTACAGAGGAGGCTCCATTTAGTATGACTTCG 539
Db 421 TACCATATCAACGATTAATTAATTTGTTACAGAGGAGGCTCCATTTAGTATGACTTCG 480
Qy 540 AATATGATGCGAGTACTTATTCAGATAGAAACAAATATTCAGAGGCTGTTGATGCTG 599
Db 481 CATATGATGCGAGTACTTATTCAGATAGAAACAAATATTCAGAGGCTGTTGATGCTG 540
Qy 600 TCTGACGCGAGTTTGGGAAATGATCAAGCAAAAGCGACCAAGTTCGCGGTGATAT 659
Db 541 TCTGACGATGATTTGGGAAATGATCAAGCAAAAGCGACCAAGTTCGCGGTGATAT 600
Qy 660 CATATCTGACAGCTGCGAATTAACACATCAAGCTGAGAGGAGTATGATATTCGAT 719
Db 601 AATATTTTAAACAGTGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
Qy 720 TTGGAGAGGAGTCTGTAAGAGGAGAAATTTGTCATTAACCATTCAGAGGCTCAAG 779
Db 661 GTAGAGAGTGAATATTCGAAAGTTGGGAGTATTTGTCATTAACCATTCAGAGGCTCAAG 720
Qy 780 GGGAGCAGTGTCTTCGATGTTTATTTATGATGCTGTAAGAAACAAATGTTATTAAT 835
Db 721 GGGAGCAGTGTCTTCGATGTTTATTTATGATGCTGTAAGAAACAAATGTTATTAAT 780
Qy 840 GGGATATTAAGGAGAGCAACCTTTTGAAGGCAAAAGAAATGAGTTTCAATTTGATTCG 899
Db 781 GAGATATTCGAGAGAGGAGCAACCTTTTGAAGGCAAAAGAAATGAGTTTCAATTTGATTCG 840
Qy 960 AATCTTATTTGATGAATTTTGAAGAGATTTACATCACTTTCACCCGAGCT 959
Db 841 AATCTTTCCTTGATGAATTTTACCAAGATCAACCAATGTTTATTAATTAATTAATTAAT 900
Qy 960 GGTATGAGTGAACATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1019
Db 961 GGGATATTCGATTTTACATTTCTAAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 960
Qy 1020 TCAAGATACATCAAGAAATTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1079
Db 961 ATTGGAAGAAC-----TCTGAAATTAACCTTTAGAGAAACAAATTAATTAATTAAGTGAAT 1014
Qy 1080 AAGGATTAAGTTCATTAATTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1139
Db 1015 AAGGATTAAGTTCATTAATTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1074
Qy 1140 AATGAGAAAGCTATTTATTTATGATCAAAAGAGAGAGTATTAAGTGAATTAAGTGAAT 1199
Db 1075 AATGAGAGAAATTTATTTTATGATCAAAAGAGAGTATTAAGTGAATTAAGTGAATTAAGTGAAT 1134
Qy 1200 AATTAACCAAGAGAGGAGTCTTTATTTTGAAGGATTAATTAAGTGAATTAAGTGAATTAAGTGAAT 1259
Db 1135 ATGATCAAGAGAGGAGGAGTCTTTATTTTGAAGGAGGATTTATTAAGTGAATTAAGTGAATTAAGTGAAT 1194

Qy 1260 AACCAACTTGGCGAGAGAGCTGGATACATGTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1319
Db 1195 AATTAAGCTGTAAGAGAGGAGGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1254
Qy 1320 GTAAATGAGGAGAGAGATGATGATCTTCTTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1379
Db 1255 GTAAATGAGGAGAGAGATGATGATCTTCTTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1314
Qy 1380 GCCAAAGGAGAAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1439
Db 1315 GCCAAAGGAGAAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1374
Qy 1440 CAGGAGAGAGATTAAGGAGCAACAAAGAGCTTTAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1499
Db 1375 CAGGAGAGAGATTAAGGAGCAACAAAGAGCTTTAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1434
Qy 1500 GGGACTGTTCAATTAAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1559
Db 1435 GGGACTGTTCAATTAAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1494
Qy 1560 CGTGTGTGTGCTTATGATCTTAAAGGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1619
Db 1495 CGTGTGTGTGCTTATGATCTTAAAGGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1554
Qy 1620 GAGAGAGGAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1679
Db 1555 GAGAGAGGAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1614
Qy 1680 GGGAGAGAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1736
Db 1615 GGGAGAGAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1674
Qy 1737 AAGAAATTTGCTTCAAGGCTTGTGTTTGGCAAAAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1796
Db 1675 AAGAAATTTGCTTCAAGGCTTGTGTTTGGCAAAAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1734
Qy 1797 TTAACCTTATTTTAAACCAACCAAGAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1856
Db 1735 TTAACCTTATTTTAAACCAACCAAGAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1794
Qy 1857 AATTTAAAGGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1916
Db 1795 AATTTAAAGGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1854
Qy 1917 CCGAGAGCTTCAATTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1976
Db 1855 CCGAGAGCTTCAATTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1914
Qy 1977 GAAATTTGTGAGGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2036
Db 1915 GAAATTTGTGAGGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 1974
Qy 2037 AAGGAGAGAGTGTGAGTGTCTTGTGAGATTTCTTCAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2096
Db 1975 AAGGAGAGAGTGTGAGTGTCTTGTGAGATTTCTTCAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2034
Qy 2097 AGCAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2156
Db 2035 AGCAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2094
Qy 2157 CTTTCAAGTGAAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2216
Db 2095 CTTTCAAGTGAAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2154
Qy 2217 ATTAATTTTAAACCAACCAAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2276
Db 2155 ATTAATTTTAAACCAACCAAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2214
Qy 2277 GGGAGTGAAGAGTGAAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2336
Db 2215 GGGAGTGAAGAGTGAAGAGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 2274

QY 2337 TTACATTAGCAACATGCGACCCAAATAGCCATATTCGACTTCCGACATTCACACT 2396
 DB 2275 TTTCATTAGCAACATGCGACCCAAATAGCCATATTCGACTTCCGACATTCACACT 2394
 QY 2397 ----- 2396
 DB 2335 GCAAGCGTAATATATGCGACTTATGCGGAGATGCAATTTAGCGGATACCTAGCCGTTT 2394
 QY 2397 -----GCA 2399
 DB 2395 ACATTAAGCAATCAACACACAGATTGGGACATGAGCTTCACTGACGACGCTCAACGA 2454
 QY 2400 ACGGTGATATATGCAAACTTGAACGGTATATGCAATTTAGCGATTCAGCTCAATTTCT 2459
 DB 2455 AAGTGAATATGCAAACTTGAACGGTATATGCAATTTAGCGATTCAGCTCAATTTCT 2514
 QY 2460 TTAATAAAACAGCAATTTTGGCAAAATTCAGGAGACAAAGGACAAACAGTGAAGCTTG 2519
 DB 2515 TTAATAAAACAGCAATTTTGGCAAAATTCAGGAGACAAAGGACAAACAGTGAAGCTTG 2574
 QY 2520 GAAATGCGACTTGGCAATGCGTACGATACCTAATTCAGATTTAAAGCTTAATAC 2579
 DB 2575 GAAATGCGACTTGGCAATGCGTACGATACCTAATTCAGATTTAAAGCTTAATAC 2634
 QY 2580 AGTACGATCAGCTTAATTTAGCTTATTCAGCTACCTAAACAAATACGCACT ---CGC 2636
 DB 2635 AGTACGATCAGCTTAAATTTAGCTTATTCAGCTACCTAAATATGCGCACTGCGC 2694
 QY 2637 GGTTCATTAGACGGAACAAACGCGACATGCGGACAACTGCTTCAACGCAATGCA 2696
 DB 2695 GGTTCATTAGACGGAACAAACGCGACATGCGGACAACTGCTTCAACGCAATGCA 2754
 QY 2697 GTAAATGCTAAATTTAGTGGGCAAGGACATTCGATTTCACTTTATTTGGCTAT 2756
 DB 2755 GTAAATGCTAAATTTAGTGGGCAAGGACATTCGATTTCACTTTATTTGGCTAT 2814
 QY 2757 AAAAGCGATTAATTTAATTTATTCATGACGCTGAGGCGCATTAATATTCGTTGCG 2816
 DB 2815 GAAAGCGATTAATTTAATTTATTCATGACGCTGAGGCGCATTAATATTCGTTGCG 2874
 QY 2817 AACACAGCGAAGAAACCGGAAACCTTGGAGCAATTAATTTGGTGAAGGCAAGATAT 2876
 DB 2875 AACACAGCGAAGAAACCGGAAACCTTGGAGCAATTAATTTGGTGAAGGCAAGATAT 2934
 QY 2877 CAACCTTATCAGATTAAGCTCAATTTACTTTAGAAATATACCAAGCTTATGCGAGTGA 2936
 DB 2935 AAACCTTATCAGAACTCAATTTACTTTAGAAATATACCAAGCTTATGCGAGTGA 2994
 QY 2937 TTAAGTTAATTTAGTGAAGATGATGCGCAATTCGCTTGCATTAACCAATTAAGAG 2996
 DB 2995 TTAAGTTAATTTAGTGAAGATGATGCGCAATTCGCTTGCATTAACCAATTAAGAG 3054
 QY 2997 CAGGAATTCACAAATGATTTAGTGAAGACAGGACAGAAACAACTTGAAGGCA 3056
 DB 3055 CAGGAATTCGCTGCTGATTTAGTGAAGACAGGACAGAAACAACTTGAAGGCA 3114
 QY 3057 CAAGTTGAACGACTGCTTAAACCAACAGAGTGAACCAAAAGCTGCTCAAGAGAGCA 3116
 DB 3115 CAAGTTGAACGACTGCTTAAACCAACAGAGTGAACCAAAAGCTGCTCAAGAGAGCA 3168
 QY 3117 GCGAGAGCAAGGCTTCTGATACCTGCTGATTAAGGCTGTTAAAGCAATTAAGAGC 3176
 DB 3169 ---AAGAGGCTGCTGATTAAGGCTGCTGCTTAAAGGCTGTTAAAGCAATTAAGAGC 3225
 QY 3177 AAACAAAGCTGAATGCTGTAACCAACAAAGATTAAGGCAAAAGAGAGAGAGAG 3236
 DB 3226 AAACAAAGCTGAATGCTGTAACCAACAAAGATTAAGGCAAAAGAGAGAGAGAG 3285
 QY 3237 TCAAAAGAGC-----AGTGTTCGATCCCTGCTTGAACAGCTGTT---C 3284
 DB 3286 TCAAAAGAGCATTGAGAGAGAGCTTCTGATACCCGCTGATCTTAAGCAGTTTAAC 3345
 QY 3285 GCATTAAGAGAGAGAGCTTGAATGATGATCCCAACAGATTCGAAAAAGATGCTTA 3344

DB 3346 GTATTAGAGCGGACCTTAAGCTTATTAATGCGCAACCGCAACCAAAAGAGAGTCA 3405
 QY 3345 GCTCA---GAAAGAGCGGAAAAACAGCGAAACAAAGAGCTTATGAGCGCTTATTC 3401
 DB 3406 GCTCAAG 3465
 QY 3402 AATAGTGCCTTATGAGATTTATCTGCAAGATTAATAGTGTCTTCTGTTCAAGATGA 3461
 DB 3466 AATAGTGCCTTATGAGATTTATCTGCAAGATTAATAGTGTCTTCTGTTCAAGATGA 3525
 QY 3462 TTAGATGCTCTTTTGTATGATCAAGCAATTCGCGGTGAGCAAAATATGCGACAGAT 3521
 DB 3526 TTGATGCTCTTTTGTATGATCAAGCAATTCGCGGTGAGCAAAATATGCGACAGAT 3585
 QY 3522 AAAAGAGCTTATGATTTGATGAGGCTTCCGCTTATATGCGAGCAAAACAACTTACG 3581
 DB 3586 AAAAGAGCTTATGATTTGATGAGGCTTCCGCTTATATGCGAGCAAAACAACTTACG 3642
 QY 3582 CAATTTGGGCTGCAAAAGAGCTTACCTAATGAGCAATTTGGGCGAGTTTCTGCAATAC 3641
 DB 3643 CAATTTGGGCTGCAAAAGAGCTTACCTAATGAGCAATTTGGGCGAGTTTCTGCAATAC 3702
 QY 3642 GGTTCAGATTAATACCTTGTATGAAACAGTAAATATGAGCGCAATTAAGATGATGCG 3701
 DB 3703 GGTTCAGATTAATACCTTGTATGAAACAGTAAATATGAGCGCAATTAAGATGATGCG 3762
 QY 3702 GGTTCGCGCAATATGCAATGGGCGATTTACATTTGGTGTAAAGGAGGCAATG 3761
 DB 3763 GGTTCGCGCAATATGCAATGGGCGATTTACATTTGGTGTAAAGGAGGCAATG 3822
 QY 3762 AGTGGAGGTAATATGCTGTAAGCAACAGCGCAAAATTCATGCAAAAGCAATTAAT 3821
 DB 3823 AGTGGAGGTAATATGCTGTAAGCAACAGCGCAAAATTCATGCAAAAGCAATTAAT 3882
 QY 3822 GCGGTGAATGCAATATGATATGCTGCTTATGAGCAATTTGGGCAATTTGGGCAAT 3881
 DB 3883 GCGGTGAATGCAATATGATATGCTGCTTATGAGCAATTTGGGCAATTTGGGCAAT 3942
 QY 3882 GTTAATGCTATTTTATGATGAGTGAATTAATCAATGAGAGAGTGAAGTGAAGAG 3941
 DB 3943 GTTAATGCTATTTTATGATGAGTGAATTAATCAATGAGAGAGTGAAGTGAAGAG 4002
 QY 3942 CCGAGCTTGTATTAATGCTTATGAGTGAATTTGCAATTTGCAATTTGCTTATGCTG 4001
 DB 4003 CCGAGCTTGTATTAATGCTTATGAGTGAATTTGCAATTTGCAATTTGCTTATGCTG 4062
 QY 4002 ACGATTAATTAACGCTTAAGCTTATTTCTGCTCAATTAATGCTTATGCTTATGCTG 4061
 DB 4063 ACGATTAATTAACGCTTATGCTTATTTCTGCTCAATTAATGCTTATGCTTATGCTG 4122
 QY 4062 AACGTAAACACAGGTAATTCAGAGTGTGCAACACCAATTTGAGAGTGAATGAGCA 4121
 DB 4123 AACGTAAACACAGTGAATTCAGAGTGTGCAACACCAATTTGAGAGTGAATGAGCA 4182
 QY 4122 AAAGAGTGGGATTAAGAGCGAAGATTTTCAATTTGCAATTTGCTTATGCTTATGCTG 4181
 DB 4183 AAAGAGTGGGATTAAGAGCGAAGATTTTCAATTTGCAATTTGCTTATGCTTATGCTG 4242
 QY 4182 TCTCAAGGTTCAACACTGCGCAACACCAAAATGAGGAGTGAATTTGGGCTGATG 4241
 DB 4243 TCTCAAGGTTCAACACTGCGTAAACACCAAAATGAGGAGTGAATTTGGGCTGATG 4302
 QY 4242 TAA 4244
 DB 4303 TAA 4305

RESULT 8
 US-08-426-787-1/C
 ; Sequence 1, Application US/08426787
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert D. Fleischmann

QY	420	TTAAAGTTGTAAGAAATTA	CTCAAA--AAAGTAATTA	CACTCTATAGAGAC	476
Db	278167	TATCAAAATGTAAAGAAATTA	TATTAACAAGCTGGGAGAAAGACATCCTTATGTGGA	278105	
QY	477	GATTACCAATATCAAGATTA	CTAATATCGTTACGAGAGCGCTCAATGATATGACT	536	
Db	278107	GATTATATATGCTCTGTTTCA	TAAATTTGTACTGAGGTGAACCTGCGGTATGACA	278048	
QY	537	TCGAATATGATGAGTACTTA	TTCAGATGAGCAAAATATCCAGACGCTGTGATC	596	
Db	278047	ACAATATGGAATGAAAGTAT	ATAGCTGATAGAGAACTATCTTGAGCGTGTACATTA	277988	
QY	597	GGCTGTGAGCGCACTTTT	GGCGAATGATCAAGCAAAAGCG-----	639	
Db	277987	GGCTGAGAGCGTACGATAT	GGCGTACAGATAAGTAGMAAGAAAGAACTACATAGTTCA	277928	
QY	640	-ACCAATGTCGGGTGCAAT	CACTATCTGAGAGCTGGCAATACACATAGCTGGA	698	
Db	277927	TATATGTCTCAGCTGCAT	ATCGTATCTTACTGAGAAATACCATCTCAGATGGA	277868	
QY	699	GCAAGTATGATATGTAT	ATTGGAGCGCATGTGTGAACGGGAAATATGCTTCA	758	
Db	277867	AATGTATATGTAACAGTCAT	CTTAGTGTATAGTATGACCTTATATATATGTGCA	277808	
QY	759	TTACGATATGACGCTCAAA	GGGAGACAGTCTTCCGATGTTATATATGATGCTGAA	818	
Db	277807	TTACCAAGCGGTGTCTTAA	GGCGATAGCGGTGGCCATGTTTATTTATAGTGCAG	277748	
QY	819	AAACAAAATGTTAATTA	TATGGAATTTACGAGAGGCAACCTTTTGAAGCAAGAA	878	
Db	277747	AAACAAATGCTTATTAAT	TATGCTGATTAACAATGCGATCTTTTGTGGAAGGT	277688	
QY	879	AATGCTTCATATGTTTGG	CAATCTATTTGAT-----GAAATTTTC	923	
Db	277687	AATGCTTTCATATATAG	TATAGCTGATTAACAATGCGATCTTTTGTGGAAGGT	277628	
QY	924	GAAAGATTAACAATACAT	TAACCCGAGCTGTATGATGATGACACATTAATG	983	
Db	277627	CCTATGTTTTCAGCGCTAT	ATTTCCCAATAATAGCAATTATCTTTGATCAAT	277568	
QY	984	GGAAATGATATGCTCAG	CGGCTATATCTCAGAAATTCAGAAATCAATGAAATTA	1043	
Db	277567	AATGATGTACAGTAAAT	TAATCTTAATCTGACCTAGTAAGATGCTTAAAGCAAA	277508	
QY	1044	ATTACGTTAGCAATATG	ATTAAGCT-----TTGAAAGGAGATTAAGTTCA	1094	
Db	277507	TCAGATGAGACGTGAA	GTATTTATTAATCAATCCGCTTAAAGAACT	277448	
QY	1095	AATCTGATATAGAGAC	CTATATTTATCTCCAGTTTAAACAATGAGAAAGCTA	1154	
Db	277447	GTTAAAGCAACAGGCTAT	ATATTTTACCAACAGATGGAATATGMAAAATAT	277388	
QY	1155	TATTTATGATCAAAAACA	AGATCATATCTTGCACTTGCAATTAACAAGGAGCG	1214	
Db	277387	TACCTGCGCAACAGAA	AAAGAACTTACAATGCAAAATATATTAATCAAGGTCT	277328	
QY	1215	GGTGCTTATTTAGCGTAT	TATTTTACATCTCA---ATTCTAACAACTGG	1271	
Db	277327	GGTGGATTATCTTGAAG	ATTTTGTGTTAAAGCAACAAATATATTAATCTGG	277268	
QY	1272	CAGAAGCTGCAATCAT	GTAAAGTAAATAGCAACCGTTACTTGGAAAGTAAATGGCTG	1331	
Db	277267	CAAGCTCAGGGGTATCT	ATATGACAAGATGCACTGTGTAATGAAATGTCAATCT	277208	
QY	1332	GAAATATGCACTTTCTAA	ATTTGTAAGAAATTTGCAATTTGCAACGTTCAACCGCAAGGAGAA	1391	
Db	277207	GAAATGATGCTTATCTA	AAATTTGTAATGAGCACTTTATTAATGATGAGGAAG	277148	
QY	1392	AATTAAGTTGCATCAG	GTAGCGATGTTAAGCATTTTGGACAGACAGCAAGAT	1451	
Db	277147	AATTTGGAAGTTTAAG	TGCGGTAAAGCAAGTCACTTAATCAACAGCAATGAA	277088	


```

Db 274948 CAAAAGCCTTAGATATATGACGATTTGGGGGGTTTCTGCAATAGCCGTTGAGATAT 274889
QY 3654 ACCCTTGATGACAGGTTTAAATATACGCGAATTAAGATATGTCGGTTTGGCCAA 3713
Db 274888 ACCCTTGACGAAAGGTTTAAATATACGCGAATTAAGATATGTCGGTTTGGCCAA 274829
QY 3714 TATCATATGCGGCGATTTTCAATTTGCTTAAAGTGGGAAAGGAAATTCAGTCCAGTAA 3773
Db 274828 TATCATATGCGGCGATTTTCAATTTGCTTAAAGTGGGAAAGGAAATTCAGTCCAGTAA 274769
QY 3774 ATGCTGAAAGAAAGGCGGAAATTCATGCAATGCAATTAATTAATTAATTAATTAATTA 3833
Db 274768 ATGCTGAAAGAAAGGCGGAAATTCATGCAATGCAATTAATTAATTAATTAATTAATTA 274709
QY 3834 AGTTATCAGTCCGTTTAAAGGCGATTTGGGCAATGAGCTTATTTTGGAGTAAATTCGAT 3893
Db 274708 AGTTATCAGTCCGTTTAAAGGCGATTTGGGCAATGAGCTTATTTTGGAGTAAATTCGAT 274649
QY 3894 TTTATTTGAAGTGAATTTATCATATCTGAGAGAGTGAAGTGAAGAGTGAAGAGTGAAG 3953
Db 274648 TTTATTTGAAGTGAATTTATCATATCTGAGAGAGTGAAGTGAAGAGTGAAGAGTGAAG 274589
QY 3954 TTTATTCGCTTATATCTGCGATTTGAGTTGATTAATTAATTAATTAATTAATTAATTA 4013
Db 274588 TTTATTCGCTTATATCTGCGATTTGAGTTGATTAATTAATTAATTAATTAATTAATTA 274529
QY 4014 AGCTTAAAGCTTATTTCTTCTGCGATTTGAGTTGATTAATTAATTAATTAATTAATTA 4073
Db 274528 AGCTTAAAGCTTATTTCTTCTGCGATTTGAGTTGATTAATTAATTAATTAATTAATTA 274469
QY 4074 AGCTTAAATCTCAAGCTTGTGACACATTTGAGCTTATTTGGCAAAAGAGTGGGA 4133
Db 274468 ACTGTAATATGACAGATTTGTCACATTTGAGCTTATTTGGCAAAAGAGTGGGA 274409
QY 4134 TTAAGGCAAAATTTTATATTTTCAATTTTCCGCTTATTAATTAATTAATTAATTAATTA 4193
Db 274408 TTAAGGCAAAATTTTATATTTTCAATTTTCCGCTTATTAATTAATTAATTAATTAATTA 274349
QY 4194 CAAGTGGCAAAAGCAAAATTTGCGGCTGAAATTTGGCTTATTTGGTGAATTAATTAATTA 4253
Db 274348 CAAGTGGCAAAAGCAAAATTTGCGGCTGAAATTTGGCTTATTTGGTGAATTAATTAATTA 274289
QY 4254 TAAATTTATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4313
Db 274288 TAAATTTATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 274234
QY 4314 AATTAAT 4319
Db 274233 AATTAAT 274228

RESULT 9
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; FILE REFERENCE: PH18692C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; PRIOR APPLICATION NUMBER: US 05/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4416)..(4416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

```


[illegible]

Dd 274828 TATCATGCGGGGAGATTTCACAACTTGTGTTAATGTGGTGCGGGAAATTAGTGGAGTTAA 274769

Qy 3774 ATGCCTGAAGAACAAGCCGAAAAATTCATCAGAAAAGCCAATAATTATGCGCTGAATSCA 3833

Dd 274768 ATGCCTGAAGAACAAGCCGAAAAATTCATCAGAAAAGCCAATAATTATGCGAGTAGATCA 274709

Qy 3834 AGTTATCAGTTCGGTTTTAGGSGCAAATGGSGCAATTCAGACTTATTTTGGAGTTATGCTAT 3893

Dd 274708 AGTATCAGTTCGGTTTTAGGSGCAAATGGSGCAATTCAGACTTATTTTGGAGTTATGCTAT 274649

Qy 3894 TTATTTGAACGTGAAAAATTATCAATCTGAGAGTAGAGTAAGTAAGCAAGCCCTGCGCA 3953

Dd 274648 TTATTTGAACGTGAAAAATTATCAATCTGAGAGTAGAGTAAGTAAGCAAGCCCTGCGCA 274589

Qy 3954 TTATATGCCCTTTATATGCTGGCAATTCGAGTTGATTTATACATTTACTCCGACAGATTATC 4013

Dd 274588 TTATATGCCCTTTATATGCTGGCAATTCGAGTTGATTTATACATTTACTCCGACAGATTATC 274529

Qy 4014 AGCGTTAGACCCTTATTTCTGTCAMATTANGTATGATTTGATTAAGCCCTAACGTACAAAC 4073

Dd 274528 ACGGTTTAGCCTTATTTCTTCTTGTCATATGTTATATTTTGAACCTTAAACGTCAAAAC 274463

Qy 4074 ACGETTAATCTCACGGTGTGCAACAACATTTGGACGTTATTGGCAAAAAGAGTGGCA 4133

Dd 274468 ACTGTAATATGACAGCACTTGCAACATCATTTTGGGCGTTATTGGCAAAAAGAGTGGCA 274409

Qy 4134 TTAAAGCGAATAATTTTCAATTCGAAATTTGCGSTTTTATCTGAAATCTCAAGGTGA 4193

Dd 274408 TTAAAGCGAATAATTTTCAATTCGAAATTTGCGSTTTTATCTGAAATCTCAAGGTGG 274349

Qy 4194 CAATCGGCAACAGCAAAAATGTGGCGTGAATTTGGCCTATGCTGGTGAATAATCAACA 4253

Dd 274348 CAACTCGGCAACAGCAAAAATGTGGCGTGAAATTTGGGGAATGTGGTGAATAATCAACA 274283

Qy 4254 TAATTATACGTTTATGTATTAACAAGGTGGCTAGATCGATCCACCTTTTATATCC 4313

Dd 274283 TAATTATACGTTTATGTATTAACAAGGTGG-----GCAATCCCACTTTTATATCC 274234

Qy 4314 ATAAT 4319
|||||

Dd 274233 ATAAT 274228

RESULT 10
US-10-329-67c-1/c
Sequence 1, Application US/10329670
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome,
Fraser Reference: PR186P1
CURRENT APPLICATION NUMBER: US/10/329,670
PRIOR FILING DATE: 2002-12-24
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,425
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1030121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURES:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..
(9921)

```
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (40810)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55309)..(55309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```

597 GGCTCTGACG3CAGTTTTCGGAATGATCAAGACAAG3GCT----- 639

Db 277987 GGTCAAGACGCTCAGCATTTGGCGTACAGATTAAGATGCAAGAAACGATGTACATAGTTCA 277992

Qy 640 -ACCAATGTCGCGTGCATATCATTA-CTGACAGCTGGCAATACACAAATCGACCGTGA 698

Db 277927 TATTATGCTCAGGTCGATATCGTTATCTTATCTGGAGSAAATCCCATCTCGAGATGGA 277866

Qy 699 GCAGGATGGAATATTCGATTTGGAGCGCATGTTGGTAAAGCGGAAATATGCTTCA 758

Db 277867 AATGTAATGTTACAGTCAAATCTTAAGTGAATATGATGTGCGCTTATCATTTATGCTTCA 277808

Qy 759 TTACGATTCGAGGCTCAAGAGGGAAGAGGTTCTCCGATGTTATTTAATGATGCTGA 818

Db 277807 TTACCAACGGGTGTTCTTAAGAGGATAGGTTGGCAATGTTATTTATGATGAGAG 277748

Qy 819 AAAAATAATGTTAATTAATGGAATATTAACGGAGGCAACCCCTTTGAAAGCAAGAA 878

Db 277747 AAGAAACAATGGCTTAATAATGCTGATTTACAAATGGGATCTTTTTCGAGAGGT 277688

Qy 879 AATGGGTTCAATGCTTCGCAATCTATTTTGAAT-----GAAATTTTC 923

Db 277887 AATGGGTTTCAATTAATGTAAGAGAAATGTTATTAATGAAGTTCTTGCGGTTGATACC 277622

Qy 924 GAAAGATTTACATACATCACTTTCAACCGAGCTGTAATAGAGTGCACAAATTAAT 983

Db 277627 CCTAGTCTTTCAAGCTATATTCCTCCATTAATAGCAATTAATCTCTTGTATCAAT 277568

Qy 984 GGAATGATATGTCACGCGGCTTAATACAGCAAAATCGAATACCATGAAATTAATA 1043

Db 277567 AATGATGTTACGAGTAATTAATCTTAATCTAGACCTTAGTAAGATGGCTTAAGGAAA 277508

Qy 1044 ATTACGTTACCAAAATATGAGTTACCT-----TTGAAAGGAAGATTAAGTTAT 1094

Db 277507 TCAAGATGGAACCTGTAAGGTAATTAATCAATCTGTAAATCAACCCGCTAAAGACAT 277448

Qy 1095 AATCCTAGATATGACAGCCATTAATTTATTTCTCAAGCTTTAAACATTTGAGAAAGCTA 1154

Db 277447 GTTAAACGACGACGAGGCTAATAATTTACACACAGAAATGGAATGGAATAATTT 277388

Qy 1155 TATTTATGATCAAAACAGATCATTAATCTTGCGATCTGCATTAACCAAGAGGCG 1214

Db 277387 TACCTTGGGACCAAGAAAGGACCTTTAACATCGAAATTAATTAATCAAGTCT 277328

Qy 1215 GATGCTCTTATTTGAGGGAATTTTACAGATCTCCA--ATTCTAACAACTTTGG 1271

Db 277327 GGTGATATACCTTTGAGGTAATTTGTTTAAAGGCAAGCAATTAATTAATCTTGG 277268

Qy 1272 CAAGAGCTGGCATACATG-TAGTAAATATGACCCGTTACTTGGAAAGTAATGGGCTG 1331

Db 277267 CAAGTGCAGGCTATCTATTGGAACAAATGCACTGTTGAATGAAAGTTTCAATCTC 277208

Qy 1332 GAAATGATCGACTTGTAAAAATGTTAAAGGAACCTTGCAAGCTTCAAGCAAAAGGGAA 1391

Db 277207 GAAATGATCTGTTATCTTAAAAATTTGATATGGAACCTTAATTAAGCTAATGTAAGGAAAG 277148

Qy 1392 AATAAGCTTCATCAGCTAGGCGATGTTAAGCTCATTTGGAGGACAGAGCAAGAT 1451

Db 277147 AATTTGGAAGTTAATGCGGGTAAAGGCAAGTCTTCTAAGTTCACAAACGACGATGA 277088

Qy 1452 CAAGGCAACAAACAGCTTTAGTAAATTTGGCTTGTATAGCGCAGAGGACTGTCAA 1511

Db 277087 GCGGCTCAAAAACAGGCTTCAAGAAAGTTGGGATTTGTAAGCGGTGAGCAACAGTTCA 277028

Qy 1512 TTAAAGGATTAACAATTTGTAACGATTAATTTATTTGGCTTTTGTTGGTGGTCTCG 1571

Db 277027 TTAAATATGTAAGATCAAGTGTATCTTAACAAATCTTAATTTTGATTTTGTTGGTGGTCC 276968

Qy 1572 TTGATCTTAACGGGCACTCATTAACCTTTAAAGCTATCCAAATCGAGATGACGGAGCA 1631

Db 276967 TTGATCTTAACGGGCACTCATTAACCTTTAGCGTATCCAAATTCGAAATGCGAAGAGGCGCG 276908

Qy 1632 ATGATTTGAACTTAATACACTCAACCGCTAATGTCACTTTATCTGGGAACGAAGC 1691


```

QY 3834 AGTATCACTTCGTTAGGCGAATGCGATTCAGCCTTATTTTGAGTATGCAAT 3893
DB 274708 AGTATCACTTCGTTAGGCGAATGCGATTCAGCCTTATTTTGAGTATGCAAT 274649
QY 3894 TTTATGAGGTAATAATATCATCTGAGAAAGTGAGGAAAGCCTAGCCTGCA 3953
DB 274648 TTTATGAGGTAATAATATCATCTGAGAAAGTGAGGAAAGCCTAGCCTGCA 274589
QY 3954 TTTATGAGGTAATAATATCATCTGAGAAAGTGAGGAAAGCCTAGCCTGCA 4013
DB 274588 TTTATGAGGTAATAATATCATCTGAGAAAGTGAGGAAAGCCTAGCCTGCA 274529
QY 4014 AGCGTTAGCCTTATTTCTGCAATATGATGTTTCAAGCGTAACGTAACAC 4073
DB 274528 AGCGTTAGCCTTATTTCTGCAATATGATGTTTCAAGCGTAACGTAACAC 274469
QY 4074 AGCGTTAGCCTTATTTCTGCAATATGATGTTTCAAGCGTAACGTAACAC 4133
DB 274468 AGCGTTAGCCTTATTTCTGCAATATGATGTTTCAAGCGTAACGTAACAC 274409
QY 4134 TTAAGGCAAGAAATTTTACATTTCCAAATTCGCTTTTAACTCAAGCTTCA 4193
DB 274408 TTAAGGCAAGAAATTTTACATTTCCAAATTCGCTTTTAACTCAAGCTTCA 274349
QY 4194 CACTCGCAAAAGCAAAATGCGGCGTAATGCGCTATGCTTGTAATAATCA 4253
DB 274348 CACTCGCAAAAGCAAAATGCGGCGTAATGCGCTATGCTTGTAATAATCA 274289
QY 4254 TAAATTTGCTTATGTAATAAGAGTGCGTCAATGAGTCCACCTTTTATTC 4313
DB 274288 TAAATTTGCTTATGTAATAAGAGTGCGTCAATGAGTCCACCTTTTATTC 274234
QY 4314 AATAAT 4319
DB 274233 AATAAT 274228

```

```

RESULT 11
US-10-329-960-1/c
Sequence 1, Application US/10329960
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
FILE REFERENCE: PB186P1
CURRENT APPLICATION NUMBER: US/10329,960
PRIOR FILING DATE: 2003-01-02
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patent version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)

```

```

LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40810)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45533)..(45533)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)

```

```

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102596)..(102596)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (123344)..(123344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (12336)..(12336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (13910)..(13910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c

```

```

FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)
OTHER INFORMATION: n equals a, t, g or c

```

Query Match 69.9%; Score 3019.8; DB 49; Length 1830121;
 Best Local Similarity 83.0%; Pred. No. 0;
 Matches 3641; Conservative 2; Mismatches 650; Indels 93; Gaps 14;

```

QY 1 TCATAGTCGTTAACTA-GTATTTTAAATACAAAATTTCTTAAATTAACATT 59
DB 278587 TCATAGTCGTTAACTAAGTATTTTAAATACAAAATTTCTTAAATTAACATT 278528

QY 60 ATGAAAAAACTGATTTTCTGTTAAATTTTAAACGCTGATTCATTAGGATAGA 119
DB 278527 ATGAAAAAACTGATTTTCTGTTAAATTTTAAACGCTGATTCATTAGGATAGA 278466

QY 120 TCGAAGCTGGGCTGTCACACTTATTTGGATGATTACAAATATATCTGATTT 179
DB 278467 TCGAAGCTGGGCTGTCACACTTATTTGGATGATTACAAATATATCTGATTT 278408

QY 180 GCCGAGATTAAGGGAAGTTTACAGTTGGGCTCAAAATATTAAGTTTAAACAACA 239
DB 278407 GCCGAGATTAAGGGAAGTTTACAGTTGGGCTCAAAATATTAAGTTTAAACAACA 278348

QY 240 GGGCAATTAGTGGCAATCATATGACAAAAGCCCGATGATTTTCTGATGTC 299
DB 278347 GGGCAATTAGTGGCAATCATATGACAAAAGCCCGATGATTTTCTGATGTC 278288

QY 300 CGTAACGGCTGGCAGGCTGTTGTAATTAATATATTTGAGGTGGACATAACGTA 359
DB 278287 CGTAACGGCTGGCAGGCTGTTGTAATTAATATATTTGAGGTGGACATAACGTA 278228

QY 360 GGAATATACAGTGTGATTTTGTGACAGGGAACAACCCGATCAATCTTTACT 419
DB 278227 GGAATATACAGTGTGATTTTGTGACAGGGAACAACCCGATCAATCTTTACT 278168

QY 420 TATAGATTTGAAGAAATTAAGTAACTAGAAA-AAAGATATTTATCTCTTATGAGAC 476
DB 278167 TATAGATTTGAAGAAATTAAGTAACTAGAAA-AAAGATATTTATCTCTTATGAGAC 278108

QY 477 GATTACCAATTAACCATTAACATTAATTCGTTACGAAGCGCTCAATGATAGCT 536
DB 278107 GATTACCAATTAACCATTAACATTAATTCGTTACGAAGCGCTCAATGATAGCT 278048

QY 537 TCGAATATGAATGCACTACTTATTCAGATAGAACAAATATTCGAAACGTTCTATTC 596
DB 278047 TCGAATATGAATGCACTACTTATTCAGATAGAACAAATATTCGAAACGTTCTATTC 277988

QY 597 GGGCTGACGCGAGTTTGGGAAATGATCAAGACAAAGCG----- 639
DB 277987 GGGCTGACGCGAGTTTGGGAAATGATCAAGACAAAGCG----- 277928

QY 640 -ACCAAGTTGCGGTGATATATTTGACAGCTGGCAATACACATACAGCTGGA 698
DB 277927 TATTAATGCTGAGGTGATATGTTATCTTACTGACGAAATACCATATACAGGTGGA 277868

```



```

Db 275707 AAGAACTGTGACCCCTTGAGCAATTAATTGATGTAAGCTTGAATTAATTAACCGTTA 275648
QY 2886 TCAAGTAACTCAAAATTTACTTTAGAAATGACACGTTGATGAGGTCATTAGCTAT 2345
Db 275647 TCGATTAAGCTCAAAATTTACTTTAGAAATGACACGTTGATGAGGTCATTAGCTAT 275588
QY 2946 AATTAGTGAAGATGATGCGAAATTCGGCTTGCAATACCCNATTAAGAGCGGAATG 3005
Db 275587 AATTAGTGAAGATGATGCGGAATTCGGCTTGCAATACCCNATTAAGAGCGGAATG 275528
QY 3006 CACATGATTTTATTAAGACGACGACGACGACGACGACGACGACGACGACGACGAC 3065
Db 275527 CACATGATTTTATTAAGACGACGACGACGACGACGACGACGACGACGACGACGAC 275468
QY 3066 CCGACTGCTAAACCAAAACAGGTAGCCCAAAAGTGGCTGTAAGAGACGACGACGAC 3125
Db 275467 CCGACTGCTAAACCAAAACAGGTAGCCCAAAAGTGGCTGTAAGAGACGACGACGAC 275417
QY 3126 GCGTTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGACGACGACGAC 3185
Db 275416 GCGTTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGACGACGACGAC 275357
QY 3186 GACCTGCTGCTGACAAACAAACAAAGTAAAGCAAAACAAACAAAGTAAAGCA 3245
Db 275356 GACCTGCTGCTGACAAACAAACAAAGTAAAGCAAAACAAACAAAGTAAAGCA 275297
QY 3246 -----GCAGTGTCTGATCCCTCTGATCAAAAGCTGTTAAAGCATTAAGACGAC 3293
Db 275296 GCAGTGTCTGATCCCTCTGATCAAAAGCTGTTAAAGCATTAAGACGACGAC 275237
QY 3294 GCGGACCTGAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3353
Db 275236 GCGGACCTGAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 275177
QY 3354 GAAAGGCAAAACAAAGCAAAACAAAGTAAAGCAAAACAAAGTAAAGCA 3413
Db 275176 G-----AACAAAGGCAAAACAAAGTAAAGCAAAACAAAGTAAAGCA 275126
QY 3414 TCAGATTTATCTCAACAGTAAATAGTATGATGATGATGATGATGATGATGATGAT 3473
Db 275125 TCAGATTTATCTCAACAGTAAATAGTATGATGATGATGATGATGATGATGATGAT 275066
QY 3474 TTGTAGATCAAGCAAAATCTGCGTGTGACAAATATGACGACGACGACGACGAC 3533
Db 275065 TTGTAGATCAAGCAAAATCTGCGTGTGACAAATATGACGACGACGACGACGAC 275006
QY 3534 GATTCGATGCGTTCGCGTCTATGACGACGACGACGACGACGACGACGACGACGAC 3593
Db 275005 GATTCGATGCGTTCGCGTCTATGACGACGACGACGACGACGACGACGACGACGAC 274949
QY 3594 CAAAAAGCTTATGATGACGATGATGATGATGATGATGATGATGATGATGATGAT 3653
Db 274948 CAAAAAGCTTATGATGACGATGATGATGATGATGATGATGATGATGATGATGAT 274889
QY 3654 ACCCTTGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 3713
Db 274888 ACCCTTGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 274829
QY 3714 TATCAATGCGGCGATTTCAATGTTGTTAAAGTGGGAAATGACGACGACGAC 3773
Db 274828 TATCAATGCGGCGATTTCAATGTTGTTAAAGTGGGAAATGACGACGACGAC 274769
QY 3774 ATGGCTGAGACAAAGCGGAAATTCATGCAAAAGCATTAATGAGCGTGAATGCA 3833
Db 274768 ATGGCTGAGACAAAGCGGAAATTCATGCAAAAGCATTAATGAGCGTGAATGCA 274709
QY 3834 AGTATACGTTCCGTTTAAAGGCAATTTGCGCTTATGTTGAGTATGATGAT 3893
Db 274708 AGTATACGTTCCGTTTAAAGGCAATTTGCGCTTATGTTGAGTATGATGAT 274649
QY 3894 TTTATGAGGATGAAATTTATCATCTGAGGAAAGGAGGAAAGGCTGACGCTGCA 3953
Db 274648 TTTATGAGGATGAAATTTATCATCTGAGGAAAGGAGGAAAGGCTGACGCTGCA 274589

```

```

QY 3954 TTTATGCGTAAATGCTGCAATTCGATGATTAATTAATTAATTAATTAATTAATTAAT 4013
Db 274588 TTTATGCGTAAATGCTGCAATTCGATGATTAATTAATTAATTAATTAATTAATTAAT 274529
QY 4014 AGCGTAAAGCTTATTTCTCGTCAATTAATGATGATGATGATGATGATGATGATGAT 4073
Db 274528 AGCGTAAAGCTTATTTCTCGTCAATTAATGATGATGATGATGATGATGATGATGAT 274469
QY 4074 ACGTAAATCTCAAGGTTGCAACACCATTTGAGAGTATTTGCAAAAGGAGGCA 4133
Db 274468 ACGTAAATCTCAAGGTTGCAACACCATTTGAGAGTATTTGCAAAAGGAGGCA 274409
QY 4134 TTAAGGCAAAATTTTACATTTCCAAATTCGCTTTATATCTCAAAATTCGAGTTCA 4193
Db 274408 TTAAGGCAAAATTTTACATTTCCAAATTCGCTTTATATCTCAAAATTCGAGTTCA 274349
QY 4194 CAACGCGCAACAGCAAAATGCGGATGAAATGCGGATGATGATGATGATGATGATGAT 4253
Db 274348 CAACGCGCAACAGCAAAATGCGGATGAAATGCGGATGATGATGATGATGATGATGAT 274289
QY 4254 TAAATTTATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4313
Db 274288 TAAATTTATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274234
QY 4314 AATTAAT 4319
Db 274233 AATTAAT 274228

RESULT 12
PCT-US03-05226-12
; Sequence 12. Application PC/TUS0305226
; GENERAL INFORMATION:
; APPLICANT: St. Geme, Joseph W.
; TITLE OF INVENTION: HEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RET/DCF/DHR
; CURRENT APPLICATION NUMBER: PCT/US03/05226
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 08/236,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 5245
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (430)..(4740)
; OTHER INFORMATION:
PCT-US03-05226-12

Query Match 53.24; Score 2990.4; DB 1; Length 5245;
Best Local Similarity 82.0%; Pred. No. 0;
Matches 3660; Conservative 0; Mismatches 636; Indels 168; Gaps 12;
QY 1 TCAATAGCTTTTACTA-GTATTTTATGCAAAATTAATTAATTAATTAATTAATTAATTAAT 59
Db 370 TCAATAGCTTTTACTA-CGATTTTATGCAAAATTAATTAATTAATTAATTAATTAATTAAT 429
QY 60 ATGAAAAAACTGATTTTCTGCTTAATTTTAAAGCTTGAATTCATTAAGGATAGTA 119
Db 430 ATGAAAAAACTGATTTTCTGCTTAATTTTAAAGCTTGAATTCATTAAGGATAGTA 489
QY 120 TCGAAGGCGGCGTGTGACACTTATTTGAGATTAATTAATTAATTAATTAATTAATTT 179
Db 490 TCGAAGGCGGCGTGTGACACTTATTTGAGATTAATTAATTAATTAATTAATTAATTT 549
QY 180 GCGGAAATTAAGGAGTCAAGTTCAGCTTGGGCTCAAAATTAATTAATTAATTAATTAATTA 239

```

Db 550 GCTGAGATTAAGGAGCTTTTCAGTTGGGGCTAAAAATATGAGTTTATATACAAAG 609
 Qy 240 GAGCAATTAGTGGACATCAATGACAAAAGCCCGATGATTTTCTGTAGTGTCA 299
 Db 610 GGGACTTTAGTGGACATCAATGACAAAAGCCCGATGATTTTCTGTAGTGTCA 669
 Qy 360 GGTAAAGGCTGGACAGCTTGGTGAATAATCAATATGTGAGCGTGGCAATACGTA 359
 Db 670 CGAAATGGGGTGGCGCATAGTAGCGGATCAGTATATGTGAGTGGCAATACGCGT 729
 Qy 360 GGAATTCAGATTTGATTTGTGTGAGGGGAAAACAACCCGATCAACATGTTTACT 419
 Db 730 GGAATTAATAGCGTTGATTTGTGAGGAGGATCCCAATCCGATCAGATCGTTTACT 789
 Qy 420 TATAGATTTGAAAAGCAATPACACAAAAGATTAATTACCTTTATGAGGAGAT 479
 Db 790 TATCAATTTGTAAGAAAGAAATATATATAGCCGCAAGATACCTTTATCAGTGTAC 849
 Qy 480 TACCAATATCCAGATTAACATTAATTTGTTAAGAGGCGCTCCAAATGATATGACTTG 539
 Db 850 TATCAGATGCTGTGTCACAAATTTGTCACTGACGCTGACACCAAGATGACAGAC 909
 Qy 540 AATATGAATGGCAGTACTTATTCAGATGAGACAAAATACCAAGATGTTGTATGGC 599
 Db 910 AATATGAATGGAAAGAACTACGCTGATTTAGTAAATATCTGATGTGTGTGTATGGT 969
 Qy 600 TCTGACGCGAGTTTGGCGAATGATCAAGACAAAGGCGACCA-----644
 Db 970 ACAGTGAACATGTTGGAGGAGCTGATGAAGAAACAAAGCAAGAGTAAAGTTCAATGG 1029
 Qy 645 GTTCCGCGTGCATTCATTTATCTGACAGCTGGCAATACACATGACGTGGAGCAGGT 704
 Db 1030 CTTCGCTGATGCTTATCTGTGAGAAATGCGAGTTACACACATTCACAAAGTGGACGGGC 1089
 Qy 705 AATGGAATTTGATTTGGAGCGCATGTTGTGAAGGGGAGAAATGTCATATCCG 764
 Db 1090 AACGCAACGATTAACCTTAAGTGAAGATACCAAAACCAATTAATGATGACCTTCTCT 1149
 Qy 765 ATTGACGCTCAAGAGGAGCAGTGTCTCCAGTGTATTTATGATCTGAAAAACA 824
 Db 1150 ACGGGTGTTCGTGTGAGATAGTGTCTCCAAATGTTTATATGATGCAATTAACAA 1209
 Qy 825 AATGCTTATTTAATGGAATTTACGGAAGGCAACCTTTGAGGCAAGAAATGGG 884
 Db 1210 AATGCTTATTTAATGGAATTTACGGAAGGCAACCTTTGAGGCAAGAAATGGG 1269
 Qy 885 TTTCATTTGCTGCAATCTTATTTTGAAGAA---TTTTCGAAGAGATTTACACCA 941
 Db 1270 TTCCATTAATTTAGAAAAATTTGTTTATGATATGTTCTTTGAGAAATTTGCTATA 1329
 Qy 942 TCACCTTACACCCGAGCTGTATGAGATGACAAATTAAGTGAATGATATGTCAG 1001
 Db 1330 ACATTTTATGAGCAGAAAGTAAAGGCTATTTATTTATTTCTTCAATTAATTAAGAACT 1389
 Qy 1002 GGGCTATATACTC-----AGAAATAGGAATCCATCAGAAATTAATTAATGCTTA 1052
 Db 1390 GGTACGCTTACCAACCAATGAAAGAGATATGCTCAATTTAAATGCAAGAGGTT 1449
 Qy 1053 GCAAAATATGATTTCTTGAAGAGAGATTAAGTCAATATCTTATGATATGAGGCA 1112
 Db 1450 CAGTATTTATGAGAGCATTAAGAAAGTAAAGACCTGTTATGCTGACAGTGTG 1509
 Qy 1113 CTAATATTTATTTCTCCAGTTTAAACATGAGAAAGCTATATTTTATGAGTCAAAA 1172
 Db 1510 GTAAATGCTTATTAACCAAGCTAATTAATGTAAGAAATTTTATCTTTGGAGTGGAGA 1569
 Qy 1173 CAGGATCATTAATCTTGGACATCTGACATTAACAAAGGCGGGTGTCTTATTTGAG 1232
 Db 1570 ACAGGAATTTAATGAAATGAAATATATATATCAAGTGTGTGTGTGATTTTGAAG 1629
 Qy 1233 GGTAAATTTAGATTTCCCAAAATCTTACCAAACTTGGCAAGAGATGGGACATGATA 1292
 Db 1630 GGTAACTTTACGGTATTTCCGAAATTAATGCACTTGGCAAGGTCGTGAGTGCATGTA 1689

Qy 1293 AGTAAAAATAGACCGTATCTTGGAAAGTAAATGGCGTGAACATGATCGACTTTCTAAA 1352
 Db 1690 GGTGAAGACAGTGTATCTTGGAAAGTAAACGGGTGGAACATGATCGCTTTCTAAA 1749
 Qy 1353 ATTGTTAAAGAACATTTGACGTTCAAGCCAAAGGAGAAATTAAGTTGATCAAGGTA 1412
 Db 1750 ATTGTTAAAGAACGTTGCAATTCACCAAAAGTGAAGAACTTAGCTCAATTAAGGTA 1809
 Qy 1413 GCGATGTAAGTCAATTTTGGAGGACGAGGACGATCAAGGCAACAAACAGCTTT 1472
 Db 1910 GTGAGCGCAAGTCAATTTTGAATCAACAGCTGATGAGACCAACCAAAACAGCTTT 1869
 Qy 1473 AGTGAATTTGCTGTGTAAGCGAGAGGAGCTTTCAATTAAGATGATTAACAAATTT 1532
 Db 1870 AAGGAATGGCATTTGATGATGAGAGCTTCCTCACTAAATGAGGAGATGAGT 1529
 Qy 1533 GATACGTAATTTATTTATTCGCTTTGTGTGTGTGCTTATGATCTTAAAGGCAATGCA 1592
 Db 1930 GATCTTAACAAATTTATTTTCGATTTGTGTGTGTGCTTATGATCTTAAAGGCAATGCA 1599
 Qy 1593 TTACCTTTAAAGTATCCAAATTCGAGAGGAGGCAATGATTTGTAACCAATATACA 1652
 Db 1990 TTACCTTTAAAGTATCCAAATTCGAGAGGAGGCAATGATTTGTAACCAATATACA 2049
 Qy 1653 ACTCAAGCGCTAATGTCATTAATCTGAGAACCAAGCAT---TGTTCTACCTAATGGA 1709
 Db 2050 ACTCAAGCGCTAATGTCATTAATCTGAGAACCAAGCATTTACTGCTCACTAATATA 2109
 Qy 1710 AATATATTAATTAATCTTGAATTAAGAAAGAAATTTGCTTACACGTTGTTGCGAA 1769
 Db 2110 AATATATTAATTAATCTTGAATTAAGAAAGAAATTTGCTTACACGTTGTTGCGAA 2169
 Qy 1770 ACAGATTAATTAATTAATCTTGAATTAAGAAAGAAATTTGCTTACACGTTGTTGCGAA 1829
 Db 2170 ACAGATTAATTAATTAATCTTGAATTAAGAAAGAAATTTGCTTACACGTTGTTGCGAA 2229
 Qy 1830 CGTACTTTGCTAATCTTCAAGTGTGTCAAATTTTAAAGGCGATTTACCAAAAGGAT 1889
 Db 2230 CGTACTTTGCTAATCTTCAAGTGTGTCAAATTTTAAAGGCGATTTACCAAAAGGAT 2289
 Qy 1890 AATCTATTTTCAAGGAGTACGACGACCGGACGCTCAATCAATTAATTAAGTGTG 1949
 Db 2290 AATCTATTTTCAAGGAGTACGACGACCGGACGCTCAATCAATTAATTAAGTGTG 2343
 Qy 1950 TCAGAAATGGAAGGATACCAAGGCGAAATTTGTGAGATCAAGATGATCAACGCT 2009
 Db 2344 CCAGAGAGCTTGGGCACTTCAAGGCGAGGTTATGATGAGATGATGATCAACGCT 2403
 Qy 2010 ACATTAATGCTGAAACTTCCAAATTTAAAGGCGAGGTCGAGTGTCTCCCAATGTT 2069
 Db 2404 ACATTAATGCTGAAACTTCCAAATTTAAAGGCGAGGTCGAGTGTCTCCCAATGTT 2463
 Qy 2070 TCTTCAATTTGAGGAATTTGACAGTCAAGATTAATGCAATGCCCATTTGTTGTG 2129
 Db 2464 TCTTCAATTTGAGGAATTTGACAGTCAAGATTAATGCAATGCCCATTTGTTGTG 2523
 Qy 2130 CCAGATTAATGGAATTTGACAGTCAAGATTTGAGGAGATTAAGGATTTGCA 2189
 Db 2524 CCAGATTAATGGAATTTGACAGTCAAGATTTGAGGAGATTAAGGATTTGCA 2583
 Qy 2190 AAGTGAATTTAACCGATTAACAAAGTTATTAATCTTAAACCAAAACCAATCAATGCT 2249
 Db 2584 AAGTGAATTTAACCGATTAACCAAGTTATTAATCTTAAACCAAAACCAATCAATGCT 2643
 Qy 2250 TCTATTAATTTAATGATATGACAGCGCAAGTAAAGTTTGAAGGTTTGAAGGTTTGAAG 2369
 Db 2644 TCTATTAATTTAATGATATGACAGCGCAAGTAAAGTTTGAAGGTTTGAAGGTTTGAAG 2703
 Qy 2310 AAGTCACTTTAAACCAATCAAGGCAATTTAATTAAGCAACATGCAACCAATTAAGG 2369
 Db 2704 AATGCACTTTAATTAATCAAGGCAATTTAATTAAGCAACATGCAACCAATTAAGG 2763

QY 2370 AATATGACCTTCCGACAACTTCAACCTGCAACGCTGATATGCAAACTTGAACGCTAT 2429
 Db 2764 AATATGCACTTTCAATACGCAATATGCAACGCTGCAACATGCAATTTGAACGCTAT 2823
 QY 2430 GTGCTTTAAGGATTTACCTCAATTTCTTTAAAAAGCAACGCTTTTGGCAACAAAT 2489
 Db 2824 GTGCTTTAAGGATTTACCTCAATTTCTTTAAAAAGCAACGCTTTTGGCAACAAAT 2883
 QY 2490 CAGGAGCAAAAGGCAAAAGCTGACGCTGGAATATGCAATGCTGACAT 2549
 Db 2884 CAGGAGCAAAAGGCAAAAGCTGACGCTGGAATATGCAATGCTGACAT 2943
 QY 2550 ACTACATGCAAGATTTAAAGCTTAATTAAGATGATGATGATGATGATGATGATGAT 2609
 Db 2944 GGCACATGCAAGATTTAAAGCTTAATTAAGATGATGATGATGATGATGATGATGAT 3003
 QY 2610 GCTACCTCAAAACATATGCAAGCT---CGCGCTTCAATTGAGAGGAAACAGCCCA 2666
 Db 3004 GCTACCTCAAAATATATGCAAGCTGACGCTGCTTCAATTGAGAGGAAACAGCCCA 3063
 QY 2667 TCGGAGCAACATGCTTCAACATTTGACATGATGATGATGATGATGATGATGATGAT 2726
 Db 3064 TCGGAGCAACATGCTTCAACATTTGACATGATGATGATGATGATGATGATGATGAT 3123
 QY 2727 TTCCATTTACTTCAATCTTTAATTGGCTAATTAAGGATTAATTAATTAATTAATTA 2786
 Db 3124 TTCCATTTACTTCAATCTTTAATTGGCTAATTAAGGATTAATTAATTAATTAATTA 3183
 QY 2787 GCTGAGGCGCATTTAATGATTTATGCTTGGCAACAGGAAAGGAAACCGGCTTGA 2846
 Db 3184 GCTGAGGCGCATTTAATGATTTATGCTTGGCAACAGGAAAGGAAACCGGCTTGA 3243
 QY 2847 CAATTAATTTGTTGTAAGCAAAAGATTAATCAACGCTTATCAATGATGATGATGAT 2906
 Db 3244 CAATTAATTTGTTGTAAGCAAAAGATTAATCAACGCTTATCAATGATGATGATGAT 3303
 QY 2907 TTGAAATTAATGCAAGCTTATGCAAGCTGATGATGATGATGATGATGATGATGAT 2966
 Db 3304 TTGAAATTAATGCAAGCTTATGCAAGCTGATGATGATGATGATGATGATGATGAT 3363
 QY 2967 GAATTCGCTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3026
 Db 3364 GAATTCGCTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3423
 QY 3027 GAGCAAGCAAGCAAGCAATTAAGCAACCAATGATGATGATGATGATGATGATGAT 3086
 Db 3424 GAGCAAGCAAGCAAGCAATTAAGCAACCAATGATGATGATGATGATGATGATGAT 3483
 QY 3087 GGTGAGCCAAAAGTGGCTCAAGAAAGCAAGC----- 3119
 Db 3484 AGTAATGCAAGAGTGGCTCAAGAAAGCAAGC----- 3143
 QY 3120 ----- 3119
 Db 3544 CAGTTAAGCTATTAAGAGCGCAAGCTGATGATGATGATGATGATGATGATGATGAT 3603
 QY 3120 -----AGAGCAAGCTTCTGATACCTGCTGATCAAGCTG 3158
 Db 3604 GCAAAAAGTGGCTCAAGAAAGAGCGCTGTTTCTGATACCTGCTGATCAAGCTG 3663
 QY 3159 TTAAGCGATTAAGCAAGCAAGCTGATGATGATGATGATGATGATGATGATGATGAT 3218
 Db 3664 TTAAGCGATTAAGCAAGCAAGCTGATGATGATGATGATGATGATGATGATGATGAT 3723
 QY 3219 AAAAGCAAAAAGTGGCTCAAGAAAGAGCGCTTCTGATACCTGCTGATCAAGCT 3278
 Db 3724 AAAAGCAAAAAGTGGCTCAAGAAAGAGCGCTTCTGATACCTGCTGATCAAGCT 3783
 QY 3279 CTGTTCT---GCATTAGAGCGCACTGATGATGATGATGATGATGATGATGATGAT 3335
 Db 3784 CGTTAAAGGATTAAGCAAGCTCAAGCTGATGATGATGATGATGATGATGATGATGAT 3843
 QY 3336 GATCTCTAGCTCAAGCAAGCGGAAAGCAAGCAAGCAAGCAAGCAAGCTGATGATGAT 3395

Db 3844 GAACTCAAGATCA-----GAGAAACAGCAAGCAAGCAAGCTGATGATGATGAT 3894
 QY 3396 TATTCAAATGATGCTTATCAAGATTAATGATGATGATGATGATGATGATGATGAT 3455
 Db 3895 TATTCAAATGATGCTTATCAAGATTAATGATGATGATGATGATGATGATGATGAT 3954
 QY 3456 GATGATTAATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3515
 Db 3955 GATGATTAATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4014
 QY 3516 CAGATTAAGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3575
 Db 4015 CAGATTAAGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4071
 QY 3576 TTAAGCTCAATTTGGGCTGCAAAAGCTTACGATGATGATGATGATGATGATGATGAT 3635
 Db 4072 TTAAGCTCAATTTGGGCTGCAAAAGCTTACGATGATGATGATGATGATGATGATGAT 4131
 QY 3636 CATAGCGCTTCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3695
 Db 4132 CATAGCGCTTCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4191
 QY 3696 ATGTCGGCTTTTCCCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3755
 Db 4192 ATGTCGGCTTTTCCCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 4251
 QY 3756 GGAATCAAGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3815
 Db 4252 GGAATCAAGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4311
 QY 3816 AATATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3875
 Db 4312 AATATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4371
 QY 3876 TTTGAGCTTAATGCTATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 3935
 Db 4372 TTTGAGCTTAATGCTATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4431
 QY 3936 AAAAGCGCTAGCTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3995
 Db 4432 AAAAGCGCTAGCTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 4491
 QY 3996 ACTCCGACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4055
 Db 4492 ACTCCGACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4511
 QY 4056 AAGCTTAAGTCAAAAGCAAGCTTAATGATGATGATGATGATGATGATGATGATGAT 4115
 Db 4512 AAGCTTAAGTCAAAAGCAAGCTTAATGATGATGATGATGATGATGATGATGATGAT 4571
 QY 4116 TGGCAAAAAGAGTGGCTTAAGGCAAGAAATTTAATCAATTCGATGATGATGATGAT 4175
 Db 4612 TGGCAAAAAGAGTGGCTTAAGGCAAGAAATTTAATCAATTCGATGATGATGATGAT 4671
 QY 4176 TCAAAAATCAAGGTTCAAACTCGCAAAAGCAAGAAATGAGGCTGATGATGATGATGAT 4235
 Db 4672 TCAAAAATCAAGGTTCAAACTCGCAAAAGCAAGAAATGAGGCTGATGATGATGATGAT 4731
 QY 4236 CGTTGTAATAATCAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4295
 Db 4732 CGTTGTAATAATCAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4788
 QY 4296 TCCACCTTTTATTCGATGAT 4319
 Db 4788 ---CCACCTTTTATTCGATGAT 4810

RESULT 13

US-10-687-046-12

; Sequence 12, Application US/10687046

; GENERAL INFORMATION:

; APPLICANT: St. Gene, Joseph W.

Query Match	69.2%;	Score 2990.4;	DB 53;	Length 5245;
Best Local Similarity	82.0%;	Pred. No. 0;		
Matches 3660;	Conservative 0;	Mismatches 636;	Indels 168;	Gaps 12

QY	1	TCATAGTGTGTACCA-GTATTTTAAATACGAAATTAAGTAATTAATTAACAT	59
Db	370	TCATAGTGTGTACCAAGGATTTTAAATACGAAATTAAGTAATTAATTAACAT	429
QY	60	ATGMAAAACCTGATTTTGTCTTAATTTTAAACCGTGCATTTCAATTAAGGATGTA	119
Db	430	ATGMAAAACCTGATTTTGTCTTAATTTTAAACCGTGCATTTCAATTAAGGATGTA	489
QY	120	TCCGAAGCGTGGCGTGGTCACTATTTTGGATTTGAATACCAATTAATTCGGATTT	179
Db	490	TCCGAAGCGTGGCGAGGTCACTATTTTGGATTTGAATACCAATTAATTCGGATTT	549
QY	180	GCCGAGATAAAGGAAAGTTCAAGTTGGGCTCAAAATATTAAAGTTTATACAAACA	239
Db	550	GCTGAGATAAAGGAAAGTTTCAGTTGGGCTCAAAATATTAAAGTTTATACAAAG	609
QY	240	GGGCATTTGTGGCAATCAATGACAAAGCCCGATGATGATTTTCTGTAGTGTCA	299
Db	610	GGGCACTTGTGGCAATCAATGACAAAGCCCGATGATGATTTTCTGTAGTGTG	669
QY	300	CGTACGGCTGTGGAGCGTGTGGTGAATATATTTGAGCGTGGCATATACGTA	359
Db	670	CGAATGGGCTGTGGAGCATTTAGTACGATCACTATTTGAGTGTGGCATATACG	729
QY	360	GGATATACGATTTGATTTTGTGTGAGAGGAAACAACCCGATCAACATCGTTTACT	419
Db	730	GGATATATACGATTTGATTTTGTGTGAGAGGATCAACATCCGATCGATCGTTTACT	789
QY	420	TATTAATATGTAAACGAAATATATCAAAAAAGATATTACATCTTATGAGAGCAT	479
Db	790	TATCAATATGTAAAGAAATATATTAACGAGGAAAGTAACTTATATAGGTGAC	849
QY	480	TACCATATCAAGATTAATTAATTCGTGTACGAAGGCGCTCCAAATTAATATGACTTG	539
Db	850	TATCAATATCAAGCTGTTTGTGCAAAATTTGTCACTGACGCTGAACCAAGAAATGACGAC	909
QY	540	AATATATATGCAATATTAATTCAGATGAGAACAAATATCAAGATGCTTGTATATGCG	599
Db	910	AATATATATGCAAGATTAATTCAGATGAGAACAAATATCAAGATGCTTGTATATGCG	969
QY	600	TCTGAGCGCAGTTTGTGGCAATGATCAAGCAAAAGCGACAA-----	644
Db	970	ACAGGTGAACATGTGTGAGGACTGATGAGAGCAAAAGCAAGGATGTAAGTTTCATGG	1022
QY	645	GTTCGCGGTGATATATTTCTGACAGCTGGCAATATACACATCAGGTGTGAGACAGT	704
Db	1030	CTTCGATATGTTATCTGTGAGAAATGACAGGTATCAACACATTCACAAAGTGTGAGCGCG	1083

QY	705	AATGCATATTCGCTATTTCGGAGACGGAGATGTCGTAAGAAGCGGAGAAATACAGTGCATTAACCG	764
Db	1030	AACGGACAGGTAACTTAAGTGAAGATATACAAAACCAATATCTGAGACCTTCTCT	1149
QY	765	ATTGCAGGCTCAAAAGGAGGACAGTGGTCTCCGATGTTTATTATGATGCTGAAAAACA	824
Db	1150	ACGGGGTGTTCGTTGGAGATAGTGGTCTCCAACTGTTATTATATGATGCAATTAACCA	1209
QY	825	AAATGCTTAATTAATGGGATATTAACGGAAAGCCACCTTTTGAAGCCAAAAGAAATGGG	884
Db	1210	AAATGCTTAATTAATGGGATATGGAAACTGGTAACCTCTTCGCGAGCTGAAATAGA	1268
QY	885	TTTCAATTCGTCGCAAACTTAATTGATGAAA---TTTTCGAAAGATTTACATACA	941
Db	1270	TTCCATTTAATTAAGAAAAATAGGTTTATGATATATGCTTTGAGAGATTTGCTTA	1328
QY	942	TCACTTTACACCGAGCTGTAAATGAGATGACAAATATGATGAAATGATATGTCAG	1001
Db	1330	ACATTTTATGAGCCAGAAAGTAAACGCTATTATCACTTCTCAATTAATATGAGACT	1389
QY	1002	GGGCTAATATC-----AGAAATCAGAAATACATCAAAATTAATAATAGCTA	1052
Db	1390	GGTAAGCTTACTCAACGATGAAAAATGATAGTCTCAATTTAAAGTCAGACGTT	1449
QY	1053	GCAAATATGATTTACCTTTGAAGAGAGATTAAGTCAATCTAGATATGACGA	1112
Db	1450	CAGTATTTAATGASAGCTTTAAAGAAAAAGATTAAGAACCTGTTATGTCGACGGTGT	1509
QY	1113	CTTAATATTTATCTCAAGTTTTAAACATGAGAAACGCTAATTTTATGATCAAAAA	1172
Db	1510	GTAATAGCTTTTAAACCAAGATAAATATATGATTAATAATTTACTTGGAGATCAGGA	1569
QY	1173	CAAGATCATTAATCTTCGACTCGACATTAACAAGGAGGCGGTGCTTTATTTAG	1232
Db	1570	ACGGAACCTTAACATTAATTAATTAATTAATCAAGGTGCTGTGTGTTATTTGAG	1629
QY	1233	GGTAATTTTACATATCTCAAAATTTCTAACAACTTGGTAAGAGCTGGCAATCATGTA	1292
Db	1630	GGTAATTTTACATATCTTCAAGAAAAATATGCAATTTGCAAGTGTGGAGTCAATGTA	1689
QY	1293	AGTGAATATACACCGTCTCTTGGAAAGTAAATGCGCGTGAACAATATGCACTTCTTAA	1352
Db	1690	GGTGAAGAGTACTGTATCTTGAAGATTAACGCGGTGAACAATATGCTTCTTAA	1749
QY	1353	ATTGTAAGGACATTTGACGTTCAAGCCAAAGGGGAAATTAAGTTTCATCAGCGTA	1412
Db	1750	ATTGTAAGGACATTTGACATTAACAAGAAAGGTGAATCTTAGGCTCATTAAGGTA	1809
QY	1413	GGGATGTTAAAGTCTTTTGGACAGCAGCAGAGCATCAGGCAACAACAAGCTTT	1472
Db	1810	GGTGAAGGCAAGTCAATTTATATCAACAAGCCGATGAGAACCAACAACAAGCTTT	1869
QY	1473	AGTGAATATGCTGTGTGACGCGACAGGGAGCTGTCAATTAACGATGATAAACATT	1532
Db	1870	AAAGAGTTGGCATATGATGATGATGAGCTCACGTTCACTAATATGTCAGATCAAGTT	1929
QY	1533	GATACCGATAAATTTATTTGGCTTTGCTGTGCTGTCTTAAATCAAGGCAATCA	1592
Db	1930	GATCCTTAACAATTTATTTGCGATTTGCGTGGTGTGCTTAACTTTAAGGACATCA	1989
QY	1593	TTTAACTTTAAAGTATCAAAATTAAGGACAGAGGGGCAATATGTAACATTAATCA	1652
Db	1990	TTTAACTTTAAAGTATCAAAATTAAGGACAGAGGGGCAATATGTAACATTAATCA	2049
QY	1653	ACTCAAGCGCTAATGTCATTAATTAAGTGAAGCAAGACAT---TGTTCATCAATAGA	1709
Db	2050	ACTCAAGTGTCTAATTTATTAATTAAGTGAAGCAAGATTTATGCTCATCTAATATA	2109
QY	1710	AAATATTTAATTAATCTGATTAACAAAAAGAAATGCTCAACAAGTGTGTTGTCGAA	1769
Db	2110	AAATATTTAATTAATCTGATTAACAAAAAGAAATGCTTCACACGTTGTTGTCGAA	2169

Db 4312 AATTATGCGTGATGCAAGTTATCAAGTCCGTTTAAAGGCAATGGGCAATTCAGCCTTAT 4311
 Qy 3876 TTGGAGTATATCGTATTTTATTTGAAGTGAATAATTAATCAATCTGAGAAAGTGAAGT 3935
 Db 4372 TTGGAGTATATCGTATTTTATTTGAAGTGAATAATTAATCAATCTGAGAAAGTGAAGT 4431
 Qy 3936 AAAAGCGCTAGCTTCAATTTATTCGCTATATATGCTGCAATGCAATGTTATATACATTT 3995
 Db 4432 AAAAGCGCTAGCTTCAATTTATTCGCTATATATGCTGCAATGCAATGTTATATACATTT 4491
 Qy 3996 ACTCGACATATATATGAGGTTTACGCTTATTTCTGCTCATTTATGTTGATTTTCA 4055
 Db 4492 ACTCGACATATATATGAGGTTTACGCTTATTTCTGCTCATTTATGTTGATTTTCA 4551
 Qy 4056 AACGCTAAGTACAAACGAGTAAATCTCAGCGTGTGCAACCAATTTGAGAGTAT 4115
 Db 4552 AACGCTAAGTACAAACGAGTAAATCTCAGCGTGTGCAACCAATTTGAGAGTAT 4611
 Qy 4116 TGGCAAAAAGAGTGGATTTAAAGCGAATAATTTTACATTTCCAAATTTCCGCTTTATTC 4175
 Db 4612 TGGCAAAAAGAGTGGATTTAAAGCGAATAATTTTACATTTCCAAATTTCCGCTTTATTC 4671
 Qy 4176 TCAAAATCTGAGTTCACAACTGCGAACAAGCAAAATGTGGCGTGAATTTGAGCTAT 4235
 Db 4672 TCAAAATCTGAGTTCGCAACTGCGAACAAGCAAAATGTGGCGTGAATTTGAGCTAT 4731
 Qy 4236 CGTTGTAATAATCAACAAATTTTATCGTTTATGATTAACAAGTGGGTGAGATCAGA 4295
 Db 4732 CGTTGTAATAATCAACAAATTTTATCGTTTATGATTAACAAGTGGGTGAGATCAGA 4788
 Qy 4296 TCCGACCTTTTATTCGAATAT 4319
 Db 4789 --CCACCTTTTATTCGAATAT 4810

RESULT 14
 PCT-US03-05226-16
 ; Sequence 16, Application PC/TUS0305226
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
 ; CURRENT APPLICATION NUMBER: PCT/US03/05226
 ; PRIOR FILING DATE: 2003-02-18
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 4828
 ; TYPE: DNA
 ; ORGANISM: Haemophilus influenzae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (313)..(4548)
 ; OTHER INFORMATION:
 PCT-US03-05226-16

Query Match 68.8%; Score 2970; DB 1; Length 4828;
 Best Local Similarity 82.3%; Pred. No. 0;
 Matches 3581; Conservative 0; Mismatches 645; Indels 124; Gaps 10;

Qy 1 TCATAGTGGTAACTA-GTATTTTAAATGAAAAATTAATCTAATTAATAAAT 59
 Db 253 TCATAGTGGTAACTA-GTATTTTAAATGAAAAATTAATCTAATTAATAAAT 312
 Qy 60 ATGAAAAAAGTATTTCTGCTAATTTTAAACCGCTGCAATTCATTAAGGATGTA 119
 Db 313 ATGAAAAAAGTATTTCTGCTAATTTTAAACCGCTGCAATTCATTAAGGATGTA 372
 Qy 120 TCGCAAGCGTGGCTGCTACACTTATTTGGAGTTGATTACCAATTAATATCGTGAATTT 179

Db 373 TCGCAAGCGTGGCTGCTACACTTATTTGGAGTTGATTACCAATTAATATCGTGAATTT 432
 Qy 180 GCCGAGATTAAGGAAAGTCAAGTGGGCTCAAAATTAAGGTTATTAAGAAACA 239
 Db 433 GCCGAGATTAAGGAAAGTCAAGTGGGCTCAAAATTAAGGTTATTAAGAAACA 492
 Qy 240 GGGCAATAGTGGCAATCAATGACAAAGCCCGATGATGATTTTCTGTAGTCA 299
 Db 493 GGAATTTAGTGGCAATCAATGACAAAGCCCGATGATGATTTTCTGTAGTCA 552
 Qy 300 GGTAGCGGTGGAGCGCTTGTGAAATCAATATATGTAAGCGTGCACATACGTA 359
 Db 553 GGAATTTAGTGGCAATCAATGACAAAGCCCGATGATGATTTTCTGTAGTCA 612
 Qy 360 GATATACGAGTGTGATTTGGTGGAGAGGAAACAACCCGATCAATCGTTTACT 419
 Db 613 GGTATACGAGTGTGATTTGGTGGAGAGGAAACAACCCGATCAATCGTTTACT 672
 Qy 420 TATAGATGTAAACGAATACTACAAATAATTAATTAATCTTATAGAGCAT 479
 Db 673 TATAGATGTAAACGAATACTTATTAATAATTAATTAATCTTATAGAGCAT 732
 Qy 480 TACCATATCCAGTATTAATTAATTTGTTACAGAGCGGCTCAATTTGATTCATTCG 539
 Db 733 TACCATATCCAGTATTAATTTGTTACAGAGCGGCTCAATTTGATTCATTCG 792
 Qy 540 AATATAGATGAGTACTTATTCAGATGAAACAAATATTCAGAGCGTGTGATCGG 599
 Db 793 GATATAGATGAGTACTTATTCAGATGAAACAAATATTCAGAGCGTGTGATCGG 852
 Qy 600 TCTGACGCGAGTTTGGCGAATGATCAAGCAAAAGCGACCAAGTTCGCGTGCAT 659
 Db 853 TCGGCGTGGCGAGTTTGGCGAATGATCAAGCAAAAGCGACCAAGTTCGCGTGCAT 912
 Qy 660 CATTAATCTGACAGTGGCAATTAACCAATAGCGTGGAGCGATATGATTTGAT 719
 Db 913 CATTAATCTGACAGTGGCAATTAACCAATAGCGTGGAGCGATATGATTTGAT 972
 Qy 720 TTGGAGAGGAGTGTGTAAGCGGAGATATGTCATTCAGATTCGAGGCTCAAG 779
 Db 973 CTAGAGGAGAGTGTGTAAGCGGAGATATGTCATTCAGATTCGAGGCTCAAG 1032
 Qy 780 GGGGACAGTGTGTCCGATGTTATTTATGATCTGAAAAAACAATAAGTTAAT 839
 Db 1033 GGGGACAGTGTGTCCGATGTTATTTATGATCTGAAAAAACAATAAGTTAAT 1092
 Qy 840 GGGATTTTACGGAAGCAACCTTTTGAAGCAAAAGAAATGCGTTCAATGTTGCG 899
 Db 1093 GGGATTTTACGGAAGCAACCTTTTGAAGCAAAAGAAATGCGTTCAATGTTGCG 1152
 Qy 900 AATCTTATTTTGAATGAAATTTTGAAGATTTTCAATATCAATCACTTACACCGAGCT 959
 Db 1153 AATCTTATTTTGAATGAAATTTTGAAGATTTTCAATATCAATCACTTACACCGAGCT 1212
 Qy 960 GGTATAGAGTGTACACAAATAGTGAATATATATGTCAGAGGCTTATTAATCTAGAAA 1019
 Db 1213 GGTATAGAGTGTACACAAATAGTGAATATATATGTCAGAGGCTTATTAATCTAGAAA 1272
 Qy 1020 TCGGAATTCATAGAAATTAATTAATGTTAGCAATATAGAGTTTACCTTGAAGAG 1079
 Db 1273 TCGGAATTCATAGAAATTAATTAATGTTAGCAATATAGAGTTTACCTTGAAGAG 1132
 Qy 1080 AAGGATTAAGTTCATATCTAGATGAGGAGCTTATATTTATTCACAGTTTAAAC 1139
 Db 1333 AAGGATTAAGTTCATATCTAGATGAGGAGCTTATATTTATTCACAGTTTAAAC 1386
 Qy 1140 AATGAGAAAGCTATATTTATGATCAAAAACAAGATCAATATCTTGCATCGAC 1199
 Db 1387 AATGAGAAAGCTATATTTATGATCAAAAACAAGATCAATATCTTGCATCGAC 1446
 Qy 1200 AATTAACGAAGGCGGAGTGTCTTATTTGAGGTTAATTTTACAGTATCTCCAAATTTCT 1259

Db 1447 ATCAACCAAGTCGCGGCGTTTGTATTTTGAAGGTAACCTTACGTAATCTTCAGAAAT 1506
 Qy 1260 AACCAACTTGGCAAGACTGGATCATATGTAAGTAAATAGCACCGTTACTTGAA 1319
 Db 1507 AATGCAACTTGGCAAGTCTGGAGTCACTGTAAGTAAAGACAGTACTGTTCTGGAA 1566
 Qy 1320 GTAAATGGCTGGAACATGATGACTTTCTAAATATGTAAGAACATTTGACCTTCA 1379
 Db 1567 GTAAATGGCTGTAATGATGCTTTCTAAATATGTAAGAACATTTGACCTTCA 1626
 Qy 1380 GCCAAAGGGAATTAAGTTCATGAGGTGAGGTGAGTAAAGTCAATTTGAGAG 1439
 Db 1627 GCCAAAGGGAATTAAGTTCATGAGGTGAGGTGAGTAAAGTCAATTTGAGAG 1686
 Qy 1440 CAGGACGATCAAGGCAACAAAGCCCTTAGTAAATTTGCTTGGTGAAGCGGAGA 1499
 Db 1687 CAGGACGATCAAGGCAACAAAGCCCTTAGTAAATTTGCTTGGTGAAGCGGAGA 1746
 Qy 1500 GGGACTGTTCAATTAAGGATGATTAACATTTGATACCATTAATTTTATTTGGCTTT 1559
 Db 1747 GGTACGTTCAATTAAGGATGATTAACATTTGATACCATTAATTTTATTTGGCTTT 1806
 Qy 1560 GGTGCTGCTTGAATCTTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAG 1619
 Db 1807 GGTGCTGCTTGAATCTTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAG 1866
 Qy 1620 GACGAGGAGCAATGATGTAACCATTAACACTCAAGCCGCTATGCTACTTACT 1679
 Db 1867 GATGAGGAGCAATGATGTAACCATTAACACTCAAGCCGCTATGCTACTTACT 1926
 Qy 1680 GGGACCAAGCTTTGTTCTACCTATGAAATATATTAATTAATCTGATTAAGAAA 1739
 Db 1927 GGGACCAAGCTTTGTTCTACCTATGAAATATATTAATTAATCTGATTAAGAAA 1977
 Qy 1740 GAAATGCTTCAAGGCTTGTGTTGGCAAGCAATTAATTAAGCAATGGGCTTCA 1799
 Db 1978 ---ATGCTTTATATGTTGTTGTTGGTGAATGATGATTAATTAATGCTGAGCTTTG 2034
 Qy 1800 AACCTTATTTATTAACCAACAGAAATGCTTGTCTTCACTTCAAGTGTGTAAT 1859
 Db 2035 AATGTTACTTACCAATCCGCTTAACAAAGATTAACACTTCTTATCAAGTGTGTAAT 2094
 Qy 1860 TTAATAAGGCTATTAACCAACAAAGTAAATTTTTCAGCGGTGAGCCGACCG 1919
 Db 2095 TTAATAAGGCTATTAACCAACAAAGTAAATTTTTCAGCGGTGAGCCGACCG 2154
 Qy 1920 CAGGCTTCAATCAATTAATTAAGGCTTGTGTAAGTGAAGTAAACCAAGGCGA 1979
 Db 2155 CAGGCTTCAATCAATTAATTAAGGCTTGTGTAAGTGAAGTAAACCAAGGCGA 2208
 Qy 1980 AATGTTGAGTCAATGATGATCAACCGTCAATTAAGCTGTAATTTCAATTA 2039
 Db 2209 GTGTTTATGATCAATGATGATCAACCGTCAATTAAGCTGTAATTTCAATTA 2268
 Qy 2040 GGGGAGTGGCGGTTTCTCGAATGTTCTGATTAAGGTAATTTGACGCTGAGC 2099
 Db 2269 GGGGAGTGGCGGTTTCTCGAATGTTCTGATTAAGGTAATTTGACGCTGAGC 2328
 Qy 2100 AATATGCAATGCCAATTTGTTGTTGTCGAATTAACCAAAATTCATTTGACGCGT 2159
 Db 2329 AATATGCAATGCCAATTTGTTGTTGTCGAATTAACCAAAATTCATTTGACGCGT 2388
 Qy 2160 TCAATTTGACGATTAACGATTTGCAAAAAGTGAATTTAACGATTAACCAAAAGTAT 2219
 Db 2389 TCAATTTGACGATTAACGATTTGCAAAAAGTGAATTTAACGATTAACCAAAAGTAT 2448
 Qy 2220 AATTTATCAAAAACCAATGAATGCTTATTAATTAATGATTAATGCAAGCGC 2279
 Db 2449 GATTTCAATCAAAAACCAATGAATGCTTATTAATTAATGATTAATGCAAGCGC 2508
 Qy 2280 AATTTAAAGTTTGAACCAATTAATGCAATGCTTCAATTAACCAATCAAGCAATTT 2339
 Db 2509 AATTTCAATGTTTGAACCAATTAATGCAATGCTTCAATTAACCAATCAAGCAATTT 2568

Qy 2340 AATTAAGCAACATGCGACCCCAATTAAGCAATTTGCACTTCCGACATTTCACTGCA 2399
 Db 2569 AATTAAGCAACATGCGACCCCAATTAAGCAATTTGCACTTCCGACATTTCACTGCA 2628
 Qy 2400 ACGGTGATTAATGCAATTTGAACCGTAAATGCTATTAACGATTTCACTGCAATTTTCT 2459
 Db 2629 ACGGTGATTAATGCAATTTGAACCGTAAATGCTATTAACGATTTCACTGCAATTTTCT 2688
 Qy 2460 TTAATAAGGCAATTTTGGCACCAATTTGAGGAGACAAAGGACAACTGTAAGCTTTG 2519
 Db 2689 TTAATAAGGCAATTTTGGCACCAATTTGAGGAGACAAAGGACAACTGTAAGCTTTG 2748
 Qy 2520 GAAATGCACTGGAACATGCTTACGATTAATTAAGGCTTCAATTTGAGGCTTCAATTTAC 2579
 Db 2749 GAAATGCACTGGAACATGCTTACGATTAATTAAGGCTTCAATTTGAGGCTTCAATTTAC 2808
 Qy 2580 AGTACGATCACTTAAATTCAGCTTATTCAGTACCTCAATTAAGGCTTCAATTTAC 2636
 Db 2809 AGTACGATCACTTAAATTCAGCTTATTCAGTACCTCAATTAAGGCTTCAATTTAC 2868
 Qy 2637 GCTTCAATTAAGGCAACAAAGGCAATTCGACGACAAATGCTTCAATTTAC 2696
 Db 2869 GCTTCAATTAAGGCAACAAAGGCAATTCGACGACAAATGCTTCAATTTAC 2928
 Qy 2697 GTAAATGTAATTTGACTGGGCAAGGCAATTCGATTAATTTCACTTCAATTTTGGCTAT 2756
 Db 2929 GTAAATGTAATTTGACTGGGCAAGGCAATTCGATTAATTTCACTTCAATTTTGGCTAT 2988
 Qy 2757 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2816
 Db 2989 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3048
 Qy 2817 AACACGCAAGAAACCCGAAACCTTGAAGCAATTAATTTGAGGCAATTAATTTGAG 2876
 Db 3049 AACACGCAAGAAACCCGAAACCTTGAAGCAATTAATTTGAGGCAATTAATTTGAG 3108
 Qy 2877 CAACGCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2936
 Db 3109 CAACGCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3168
 Qy 2937 TTAAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2996
 Db 3169 TTAAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3228
 Qy 2997 CAGGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3056
 Db 3229 CAGGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3288
 Qy 3057 CAAGTTGACGATGCTTGAACCAACAAAGGTAAGGCTTGAAGGCA 3116
 Db 3289 CAAGTTGACGATGCTTGAACCAACAAAGGTAAGGCTTGAAGGCA 3342
 Qy 3117 GCGAGGACGCTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3176
 Db 3343 ---AGAGGCGGTTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3399
 Qy 3177 AAACAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3231
 Db 3400 AAACAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3459
 Qy 3332 TGGGCTCAAAAAGGCA---GTGTTTCTGATCCCTGCTGATCAAGGCTGTTGCA 3287
 Db 3460 TGGGCTCAAAAAGGCA---GTGTTTCTGATCCCTGCTGATCAAGGCTGTTGCA 3519
 Qy 3288 TTAAGGCGGCACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3321
 Db 3520 TTAAGGCGGCACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3579
 Qy 3322 ---AGCATGCGAA 3332
 Db 3580 GCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3639

RESULT 15
US-10-687-046-16
Sequence 16, Application US/10687046
GENERAL INFORMATION:
APPLICANT: St. Gene, Joseph W.
TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEIN
FILE REFERENCE: A-59941-1/RT/DCE/DHR
CURRENT APPLICATION NUMBER: US/10/687, 046

Query Match	68.8%	Score 2970;	DB 53;	Length 4828;
Best Local Similarity	82.3%;	Pred. No. 0;		
Matches 3581; Conservative	0;	Mismatches 645;	Indels 124;	Gaps 10;

QY	1	TCATAGTGGTTACGA-GATATTTTAAACGAAATTTCTAATTAATTAACCT	59
Dp	253	TCATAGTGGTTACGACGATATTTTAAACGAAATTTCTAATTAATTAACCT	312
QY	60	ATGAAAAAACTGATTTTCCTTAATTTTAAACCGCTTGATTCATTAAGGATAGTA	115
Dp	313	ATGAAAAAACTGATTTTCCTTAATTTTAAACCGCTTGATTCATTAAGGATAGTA	372
QY	120	TCCAGAGGTGGGTGGTCAACATTAATTTGGATTTGATTCGAAATTAATGCTGATTT	179
Dp	373	TCCAGAGGTGGGTGGTCAACATTAATTTGGATTTGATTCGAAATTAATGCTGATTT	432
QY	180	GCCGAGATTAAGGGAAGTTTCACATTTGCGGCTCAAAATATTAAGGTTTAAACAAACA	239
Dp	433	GCCGAGATTAAGGGAAGTTTCACATTTGCGGCTCAAAATATTAAGGTTTAAACAAACA	492
QY	240	GGGCAATTAGTTGGACATCATATGACAAAGCCCGCATGATTAATTTTCTGTAGTCA	299
Dp	493	GGAAATTTAGTTGGACATCATATGACAAAGCCCGCATGATTAATTTTCTGTAGTCA	552
QY	300	CGTAAACGGGTGGAGCGTGGTTGAAATTCATATTTGGAGCGTGGACATTAAGTA	359
Dp	553	CGAAATGGGGTGGCGGATTTGGTGGCGCATCATGATTAATTTGGATGGGCACTATGTA	612
QY	360	GGATATACAGATGTTATTTTGGTGGACAGAGGAAACAACCCGCATCAACATGTTTTACT	419
Dp	613	GGCTATACCATATGGATTTTGGTGGTGGAGGCAAAAATCCGATCAACATGTTTTACT	672
QY	420	TATAGATTTGAAACGAAATATACACAAAAGATTAATTTACATCTTTATGAGCAT	479
Dp	673	TATAGATTTGAAACGAAATATATTAATAAAGATCAACGATCTTTATGAGGAATC	732
QY	480	TACCAATATCCAGATTAACATTAATTTGGTTACAGAGGCGCTCCATTTGATGACTCG	539
Dp	733	TACCAACATCCAGCTTACATTAATTTGGTTACAGAGGCGCTCCATTTGATGACTCT	792
QY	540	AATATGAATGGCAGTACTTATTCAGATTAAGCAAAATATCCAAACGTTGTTGTAATGGGC	599
Dp	793	GATATGAATGGCAACAAATATTCAGATTAAGCAAAATATCCCAACCGTGGTATCGGC	852
QY	600	TCTGACGCGAGTTTGGGGAATGATCAAGACAAAGGCAACAGTTCGCGTGCATAT	659
Dp	853	TCCGGGTGGCGTTTGGGGAATGATCAAGACAAAGGCAACAGTTCGCGTGCATAT	912
QY	660	CATTATCTGACAGCTGGCAATACACACATCCAGCTGAGAGCGGTAAATGATATTCGAT	719
Dp	913	CATTATCTGACAGCGGCAATACACACATCCAGCGGAGACAGGAGCGGCTGGTCAAGT	972
QY	720	TTGGAGGCGCATTTCTGTAACCGGAGAAATATGATTCATTAACGATTTGACGCTCAAG	779

QY 2937 TTACGTTAATTAATTAAGATGATGCGAATTCGGTTCATACCATTAATAAG 2996
 Db 3169 TGGCGTTAATTAATTAAGATGATGCGAATTCGGTTCATACCATTAATAAG 3228
 QY 2997 CAGGAATTGACAAATGATTAAGTAAAGACGACAAAGACAAATTAAGACCAA 3056
 Db 3229 CAGGAATTGCGCAATTAATTAAGTAAAGACGACAAAGACAAATTAAGACCAA 3288
 QY 3057 CAGGTAACCGAATGCTTAATAACA CAAACGGTGAGCCAAAAGTGGCTCAAGAAAGCA 3116
 Db 3289 CAGGTAACCGAATGCTTAATAACA CAAACGGTGAGCCAAAAGTGGCTCAAGAAAGCA 3342
 QY 3117 GCGAGACAGCGTTTCTGATATCCCTCCCTGATCAAGCTCTTTAAACGATTAAGACC 3176
 Db 3343 ---AGAGCGCGCTTTCTGATATCCCTCCCTGATCAAGCTCTTTAAACGATTAAGACC 3399
 QY 3177 AAACAGCTGACCTGATGTAACCA---AAAGTAAAGCAAAACAAATAAG 3231
 Db 3400 GAACTCGACAGATTAATGCGCAACGACAAAGTGGCAACGCGTCAAAATCGAAGTA 3459
 QY 3232 TCGCGTAAAGAGCA---GTGTTTCTGATCCCTGCTGATCAAGCTCTTCCGA 3287
 Db 3460 ACTGCACTTAACCAAAAGACAGCAAGTTAAACCACTCAAGTAAAGCAATTTAGTC 3519
 QY 3288 TTAAAGCCGCACTTGAGGTTATGATGCCCCC----- 3321
 Db 3520 TTGCAACTGCAATGATGTAAGAAAGAAACCGCTCAGATTGTTGTAATCAAAATTA 3579
 QY 3322 -----AGCAATCGGA 3332
 Db 3580 GCTCAGTTGAATTTAAACAAACCTAAGAAAGCTTGAAGTGGTGAAGCAAGAA 3639
 QY 3333 AAAGATGCTAGCTCAAGAAAGAAACGAAACCAACCAAAAGACCTTGATCAGC 3392
 Db 3640 AAAGAGGTAAAGCTCAAGAAAGAAACGAAACCAACCAAAAGACCTTGATCAGC 3699
 QY 3393 CGTATTCAAATAGTGGCTATCAGATTAATCTGCAACAAATTAAGTACTTCTGTT 3452
 Db 3700 CGTATTCAAATAGTGGCTATCAGATTAATCTGCAACAAATTAAGTACTTCTGTT 3759
 QY 3453 CAAGATTAATGATGCTTTTGTGATCAAGCAATCTGCGTGGTGAACAAATTC 3512
 Db 3760 CAAGATTAATGATGCTTTTGTGATCAAGCAATCTGCGTGGTGAACAAATTC 3819
 QY 3513 GCACAGATTAAGAGCTATGATTCGATTCGCTTCGCTTATCAGCAGCAAGAAACG 3572
 Db 3820 TCACAGATTAAGAGCTATGATTCGATTCGCTTCGCTTATCAGCAGCAAGAAACG 3876
 QY 3573 AACTTACGTAAATGCGGTGCAAAAGCTTAAGCTAATGGAAGAAATGCGGAGTTTC 3632
 Db 3877 AACTTACGTAAATGCGGTGCAAAAGCTTAAGCTAATGGAAGAAATGCGGAGTTTC 3936
 QY 3633 TCGCATGCGCTTCAGATTAATCTTGTGATGAACAGTTAAATCAACGCGACATTAAGC 3692
 Db 3937 TCGCATGCGCTTCAGATTAATCTTGTGATGAACAGTTAAATCAACGCGACATTAAGC 3996
 QY 3693 ATGATGTCGGGTTTGGCCAAATATCAATGCGGCGATTTACAAATTTGCTTAAACGTGGA 3752
 Db 3997 ATGATGTCGGGTTTGGCCAAATATCAATGCGGCGATTTACAAATTTGCTTAAACGTGGA 4056
 QY 3753 ACGGGAATCAAGTCAAGTAAATGCGTGAAGAAACAAAGCGAAATTAATCAAGAAACG 3812
 Db 4057 ACGGGAATTAAGTCAAGTAAATGCGTGAAGAAACAAAGCGAAATTAATCAAGAAACG 4116
 QY 3813 ATAAATTAAGTCAAGTAAATGCGTGAAGAAACAAAGCGAAATTAATCAAGTCAAGTCAAG 3872
 Db 4117 ATAAATTAAGTCAAGTAAATGCGTGAAGAAACAAAGCGAAATTAATCAAGTCAAGTCAAG 4176
 QY 3873 TATTTGAGGTTAATGCTATTTATTTGAAAGTAAATTAATCAATCTGAGAAAGTGA 3932
 Db 4177 TATTTGAGGTTAATGCTATTTATTTGAAAGTAAATTAATCAATCTGAGAAAGTGA 4236
 QY 3933 GTGAAACGCTAGCTTGAATTAATGCTAATGCTGCAATGCAATGATTAATCA 3992

Db 4237 GTGAAACGCTAGCTTGAATTAATGCTAATGCTGAGTACGCGTCAATTAAGC 4296
 QY 3993 TTACTCGACAGATTAATCAAGCTTAAGCTTATTCGTCATTAATGATTAAGT 4052
 Db 4297 TTACTCGACAGATTAATCAAGCTTAAAGCTTATTCGTCATTAATGATTAAGT 4356
 QY 4053 TCAAAGCTAAGTAAACCAAGTAAATCAAGCTTAAAGCTTAAAGCTTAAAGCT 4112
 Db 4357 TCAAAGCTAAGTAAACCAAGTAAATCAAGCTTAAAGCTTAAAGCTTAAAGCT 4416
 QY 4113 TATTCGCAAAAGAGTGGGATTAAGGCGAAATTTACATTCGAAATTCGCTTT 4172
 Db 4417 TATTCGCAAAAGAGTGGGATTAAGGCGAAATTTACATTCGAAATTCGCTTT 4476
 QY 4173 ATTCGCAAAATCTCAAGGTTCAAACTCGGCAACAGCAAAATGCGGCTAAATGCGC 4232
 Db 4477 ATTCGCAAAATCTCAAGGTTCAAACTCGGTAACAGCAAAATGCGGCTAAATGCGC 4536
 QY 4233 TATCGTTGGTAAATCAATTAATTTAT 4262
 Db 4537 TATCGTTGGTAAATCAATTAATTTAT 4566

Search completed: June 24, 2004, 19:40:30
 Job time : 13149 secs

Db 5799 GGGAGGTGGTCGGCAAAATGACAAAGCTCCGATGATTTTGGATCTGTGCT 5740
Qy 300 CGTAACGGCGTGGACCTTGGTTGAAATCATATATATGAGGTGGCAATACGTA 359
Db 5735 CGTATGCTGATGCTGATTAATGAGCATCATATATTTGATGTTGCTCACTAGG 5680
Qy 360 GGAATATACAGATGTTAATTTGTCAGAGGGAACACCCGATCAACTGTTTACT 419
Db 5679 GGAATACATATGCTGATTTGTCGGAAGGACAAATCCGATCAACTGTTTACT 5620
Qy 420 TTATAGTGTGTAAGAAATTAATCAAAAAAGATATTTACATCTTTGAGCAT 479
Db 5619 TATCAATATGTTAAAGTATATATTAACCTGATCAAGCACTATATATAGGAT 5560
Qy 480 TACCAATCCAGATTAATCAATATGTTATAGAGAGGCTCCAAATGATATGACTGG 539
Db 5559 TACCAATGCTGCTTATCAATATTTGATCACTGAAAGTCTCTCTTGTATGACAGA 5500
Qy 540 AATATGAATGCGATCTTATTCAGATAGAACAAATATCCAGAACGTGCTATCCGC 599
Db 5499 AGTATGAAGGAGACAGCTATATSCAGAACAAAGAAATACCCGATGCTGTGGATGGT 5440
Qy 600 TCTGACGCGAGTTTGGCGAAATGATCAAGAAAGCGACAAAGTTCGCTGAT 659
Db 5439 TCTGATGATCAATTTGGAAATGAAACAGATTAAGAACTCAAGTACGATCACTAT 5380
Qy 660 CATATCTGACAGCTGCGATATCACACATCACTGAGAGGATATGATTCGAT 719
Db 5379 GATATCTTATGTCGCGGAATACCCATGTCGAAACGCTGACGAAATGAGATTTGAT 5320
Qy 720 TTGGAGGCGATTTTGTAAACCGGAGAAATATGTCATTAACGATGCGCTCAAG 779
Db 5319 TTAAGTGGCATGTATGTAAGAAATGATTAAGACCTATGACGAGAGATCATTT 5260
Qy 780 GGGACAGTGGTCTCCGATGTTATTTATGATGCTGAAAAACAAATATGATATAT 839
Db 5259 GAGACAGTGGTCTGATGTTATTTATGATGCTGACCGAAAAATATGCTTATAT 5200
Qy 840 GGAATTTACGAGAAAGCAACCTTTGAGGAAAGAAATGAGTTTCAATGGTTCG 899
Db 5199 GGTGATTTGGTCTGAGAAACCATCTTGTGAGAGATATCTTTGAGTATGCTCT 5140
Qy 900 AATCTTATTTGATGAAATTTTGAAGAGAT-TTAATATCATCACTTTACACCGAGC 958
Db 5139 AATATTTATGATGAAATTTTCTGCTATATCTTAAAGACATATGATTTGCA 5080
Qy 959 TGTATATGAGGTATACATTTAGTGAA-----ATGATATGCTGAGGCTCTATA 1010
Db 5079 ACTCTTAATATCAAAATGACACTACATTTCAATCAAAAAATGATGAGTACGC 5020
Qy 1011 ACTCAGAAATCAGAAATACATCAAGAAATTAATTAAGTATGAAATGAGTTTACT 1070
Db 5019 ACTATATGATGAGAGCTGTAATATTAAGTCAAACTGATATTAATATATCT 4960
Qy 1071 TTGAAGAGAGATTAAGTCTATATCTAGATATGACGACCTTAATATTTTCTGCA 1130
Db 4959 ATGAAAGGACGAGAGCTGCTATACT-----ATGATGCTGTAACATATATCTCC 4906
Qy 1131 GGTAAACAAATGAGAAACGCTATATTTATGAGATCAAAACAGAGATCATATCTTC 1190
Db 4905 GATTAATGTCAGGCAAAACCTTATTTATGATTAAGAAAGGAGTGCATCTTT 4846
Qy 1191 GCATCTGACATTAACCAAGGCGGATGCTTTATTTTGGGTAATTTTACATATCT 1250
Db 4845 ACTATATATATGATCAAGGTGCAAGGCTTTGATTTTGGAGGATTTTATATGTTG 4786
Qy 1251 CCAATTTCTAACAACTTTGCAAGAGCTGAGATATCATATGTAATTAATGACGCTT 1310
Db 4785 CCAATCATATGAAACGTAAGGAGGAGGAGGAGATATCTATCACTGATGAAGACCTTA 4726
Qy 1311 ACTGAAAGTAAATGAGTGAACATGATCACTTTCAAAATTTGGTAAAGGACATG 1370

Db 4725 ATTGGAATGTCATGCGTGAATATGATGCTTTCTAAATTTGCTAAAGGACGTTA 4666
Qy 1371 CAGCTTCAAGCCAAAGGAGGAAATTAAGGTGATCAAGGCTAGGCAATGATCATTT 1430
Db 4665 CATTTTCAAGCAAAAGGTAATTAAGCTTATGCTCAATGAGTGGCAATGATATAT 4606
Qy 1431 TTGAGACAGAGGACAGATCAAGGCAAAACAGCTTTAGTAAATGCTTGGT 1490
Db 4605 TTGAACAAAGAGAGATCAATTAATAAACAACCTTTAGTAAATGATGCTT 4546
Qy 1491 AGGCGAGAGGACCTGCTCAATTAACGATGTAACAAATTTGATCCGATTAATTTAT 1550
Db 4545 AGGCGAGAGGACCTGCTCAATTAACGATGTAACAAATTTGATCCGATTAATTTAT 4486
Qy 1551 TTGGCTTCTGAGTGGCTTATGATCTTAAGGAGCACTTAACTTTAAGCTATC 1610
Db 4485 TTGGCTTCTGAGTGGCTTATGATCTTAAGGAGCACTTAACTTTAAGCTATC 4426
Qy 1611 CAAATATCGAGAGAGGAGCATGATGTAACCATATATCAACTCAAGCGCTATATGC 1670
Db 4425 CAAATATCGAGAGAGGAGCATGATGTAACCATATATCAACTCAAGCTATATGC 4366
Qy 1671 ACTATTAATGAGAGCAAGCATTTGCTACTTATGGAATATTTAATACTTGAT 1730
Db 4365 ACTATTAATGAGAGCAAGCATTTGCTACTTATGGAATATTTAATACTTGAT 4306
Qy 1731 TACAGAAAGAAATTTGCTCAACGCTTGTGTTGGCAACAGATTAATTAACATAT 1790
Db 4305 TACAGAAAGAAATTTGCTCAACGCTTGTGTTGGCAACAGATTAATTAACATAT 4246
Qy 1791 GAGGATTAACCTTATTTAATTAACCAACAGAGATGTAATTTGCTATTCAGGT 1850
Db 4245 GAGGATTAACCTTATTTAATTAACCAACAGAGATGTAATTTGCTATTCAGGT 4186
Qy 1851 GATTAATTTAAGGCAATATTAACCAACAAAGATTAATTTTACAGGCTATG 1910
Db 4185 GATTAATTTAAGGCAATATTAACCAACAAAGATTAATTTTACAGGCTATG 4126
Qy 1911 CCGACAGCGCAAGCTCATCTTAATTAATTAAGTGTGCGAAATGAGAGTATCA 1970
Db 4125 CCGACAGCGCAAGCTCATCTTAATTAATTAAGTGTGCGAAATGAGAGTATCA 4072
Qy 1971 CAGGCAATTTGCTGAGATCAAGTGTGCAATCACTTAAAGCTGAAACCTTC 2030
Db 4071 CAGGCAATTTGCTGAGATCAAGTGTGCAATCACTTAAAGCTGAAACCTTC 4012
Qy 2031 CAAATTAAGGCGAATGCGGCTTCTGCAATCTTCTCAATTAAGGAGATG 2090
Db 4012 CAAATTAAGGCGAATGCGGCTTCTGCAATCTTCTCAATTAAGGAGATG 3952
Qy 2091 ACAGTCAAGATTAATGCAATGCAATGCAATGCTTGTGCAATCAACAAATCACTT 2150
Db 3951 ACAGTCAAGATTAATGCAATGCAATGCAATGCTTGTGCAATCAACAAATCACTT 3892
Qy 2151 TGAAGCGTCAATGAGAGATTAACGATTTGCAAAAGGAGATTAACGATCA 2210
Db 3891 TGAAGCGTCAATGAGAGATTAACGATTTGCAAAAGGAGATTAACGATCA 3832
Qy 2211 AAGATTAATTTCTATACCAAAACAGATTAATGCTTATTAATTAATGATAT 2270
Db 3831 AAGATTAATTTCTATACCAAAACAGATTAATGCTTATTAATTAATGATAT 3772
Qy 2271 GCAAGCGCAATGTAAGGTTAGCAAACTTATGCAATGCACTTAAACAAATCAC 2330
Db 3771 GCAAGCGCAATGTAAGGTTAGCAAACTTATGCAATGCACTTAAACAAATCAC 3712
Qy 2331 AGCAATTTAATTAAGCAATGCAACCAATAGGCAATTTGCACTTCCGATAT 2390
Db 3711 AGCAATTTAATTAAGCAATGCAACCAATAGGCAATTTGCACTTCCGATAT 3652
Qy 2391 TCAAGTCAAGGAGGATTAAGCAATGCAATGCAATGCAATTTAAGGATGCACT 2450
Db 3651 TCAAGTCAAGGAGGATTAAGCAATGCAATGCAATGCAATTTAAGGATGCACT 3592

QY 2451 CAATTTCTTTAAAAAAGCCATTTTTCGACCAAAATTCAGGAGACAAAGGCAACA 2510
 Db 3591 CAATTTATTTAAAAAAGCCATCTTTTCGACCAAAATTCAGGAGACAAAGGCAACA 3532
 QY 2511 GTGACGTTGGAAAAATGCGACTTGACAAATGCTTACGATCTTACATTTGGAATTTAACG 2570
 Db 3531 GTGACGTTGGAAAAATGCGACTTGACAAATGCTTACGATCTTACATTTGGAATTTAACG 3472
 QY 2571 CTAAATACAGTACGATCAGGTTAAATTCAGCTTATTCAGTACGCTCAAAATTCAGCA 2630
 Db 3471 CTAAATACAGTACGATCAGGTTAAATTCAGCTTATTCAGTACGCTCAAAATTCAGCA 3412
 QY 2631 -----GTGCGCTTTCATTTAGACGCGAAACCAACGCAATTCGCGACAAATGCTTC 2684
 Db 3411 GCGCGCTTTCGCTTTCATTTAGACGCGAAACCAACGCAATTCGCGACAAATGCTTC 3352
 QY 2685 AACCATTTGACAGTAAATGTTAAATTTAGTGGGCAAGGCAATTCATTTTCAATTCATCT 2744
 Db 3351 AACCATTTGACAGTAAATGTTAAATTTAGTGGGCAAGGCAATTCATTTTCAATTCATCT 3292
 QY 2745 TTATTTGCTATTAAGGATTAATTAATTTATTCATTAAGTGGGCGATTTACATA 2804
 Db 3231 TTATTTGCTATTAAGGATTAATTAATTTATTCATTAAGTGGGCGATTTACATA 3232
 QY 2805 TTATCTGTTGCAACACAGGCAAGAACCCGAAACCTTACGCAATTTACTTTGGTTGAA 2864
 Db 3231 TTATCTGTTGCAACACAGGCAAGAACCCGAAACCTTACGCAATTTACTTTGGTTGAA 3172
 QY 2865 AGCAAGATTAATCAACGCTTATTCAGATTAAGCTCAATTTACTTTGAAATATCAACGCTT 2924
 Db 3171 AGCAAGATTAATCAACGCTTATTCAGATTAAGCTCAATTTACTTTGAAATATCAACGCTT 3112
 QY 2925 GATGCGAGTACATTAAGTAAATTTAGTAAAGATTAAGTGGCAATTCGCTTGGATTAAC 2984
 Db 3111 GATGCGAGTACATTAAGTAAATTTAGTAAAGATTAAGTGGCAATTCGCTTGGATTAAC 3052
 QY 2985 CCAATTAAGAGAGAGAAATTCACATTAATTTAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 3044
 Db 3051 CCAATTAAGAGAGAGAAATTCACATTAATTTAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 2992
 QY 3045 TTAGAGCCCAACAAAGTTGAAACGCTGCTTAAACCAACAAAGGTTAGCCAAATGCTGGG 3104
 Db 2991 TTAGAGCCCAACAAAGTTGAAACGCTGCTTAAACCAACAAAGGTTAGCCAAATGCTGGG 2932
 QY 3105 TCAGAGAGAGAGAGAGAGAGGCTTCCGATTAACCTGCTGATTAAGAGAGAGAGAGAGAG 3164
 Db 2931 TTAGAGAGAGAGAGAGAGAGGCTTCCGATTAACCTGCTGATTAAGAGAGAGAGAGAGAG 2881
 QY 3165 GCATTTAG 3224
 Db 2881 GCATTTAG 2821
 QY 3225 AAAAAAGTGGCTCAAAAAAGCA-----GGTCTTTGATCCCTGCTGATTAAC 3275
 Db 2820 AAAAAAGTGGCTCAAAAAAGCA-----GGTCTTTGATCCCTGCTGATTAAC 2761
 QY 3276 AGCTGCTT---GCATTTAG 3332
 Db 2760 AGCTGCTT---GCATTTAG 2701
 QY 3333 AAAGTCTGCTAGCTCAAGAAAG 3392
 Db 2700 AAAGTCTGCTAGCTCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2650
 QY 3393 CGTATTTCAAAATGAGCTTATTCAGATTTATTCGCAACAAAGTAAATGTTGCTTCTGCTT 3452
 Db 2649 CGTATTTCAAAATGAGCTTATTCAGATTTATTCGCAACAAAGTAAATGTTGCTTCTGCTT 2590
 QY 3453 CAAGATTAATTAAGATCTTTTGTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 3512
 Db 2589 CAAGATTAATTAAGATCTTTTGTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2530

QY 3513 GCACAGATTAAGAGAGCTATGATTTGATGCTGCTGCTTATTCAGAGAGAGAGAGAGAGAG 3572
 Db 2529 GCACAGATTAAGAGAGCTATGATTTGATGCTGCTGCTTATTCAGAGAGAGAGAGAGAGAG 2473
 QY 3573 AACTTACGTTAAATTTGGGGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3632
 Db 2472 AACTTACGTTAAATTTGGGGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2413
 QY 3633 TCGCATAGCGCTGATTAATTTAGCTTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3692
 Db 2412 TCGCATAGCGCTGATTAATTTAGCTTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2353
 QY 3693 ATGATGCTGGGTTTTCGCAATTAATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3752
 Db 2232 ATGATGCTGGGTTTTCGCAATTAATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2293
 QY 3753 AGCGGATTCAGTGGAGTAAATTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3812
 Db 2292 AGCGGATTCAGTGGAGTAAATTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2233
 QY 3813 ATAAATTTAGCGGATTAATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3872
 Db 2232 ATAAATTTAGCGGATTAATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2173
 QY 3873 TATTTGAGATTAATTCAGATGAG 3932
 Db 2172 TATTTGAGATTAATTCAGATGAG 2113
 QY 3933 GTGAG 3992
 Db 2112 GTGAG 2053
 QY 3993 TTTATCCGAGATTAATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4052
 Db 2052 TTTATCCGAGATTAATTCAGATGAG 1993
 QY 4053 TCAAGCGTAAAGTACAAACAG 4112
 Db 1992 TCAAGCGTAAAGTACAAACAG 1933
 QY 4113 TATTTGAG 4172
 Db 1932 TATTTGAG 1873
 QY 4173 ATCTGAATTCAG 4232
 Db 1872 ATCTGAATTCAG 1813
 QY 4233 TATGTTGAG 4292
 Db 1812 TATGTTGAG 1753
 QY 4293 AGATCCAG 4319
 Db 1752 AGATCCAG 1726

RESULT 2
 PCT-US04-07001-684/C
 ; Sequence 684, Application PC/TUS0407001
 ; GENERAL INFORMATION:
 ; APPLICANT: BAYALLETZ et al.
 ; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
 ; FILE REFERENCE: 2835/38915A
 ; CURRENT APPLICATION NUMBER: PCT/US04/07001
 ; CURRENT FILING DATE: 2004-03-05
 ; PRIOR APPLICATION NUMBER: US 60/453,134
 ; NUMBER OF SEQ ID NOS: 771
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 684
 ; LENGTH: 417576
 ; TYPE: DNA

Query Match	73.9%;	Score 3193.4;	DB 1;	Length 417576;
Best Local Similarity	85.3%;	Pred. No. 0;		
Matches 3710; Conservative	0;	Mismatches 576;	Indels 61;	Gaps 11;

Dh	211735	TTAAGTGGGACATGTCGTTAAAGAAATGATTAATGACATTAATGCGACAGGAGATCATTT	211736
Qy	780	GGGAGACAGTGTCTTCGATGTTATTATATGACTCGGAAAACAAAATGCTTAATTAAT	839
Dh	211735	GGAGACAGTGGGTGCCTATGTTATTATATGATCTACGAGAAAATATGGCTTAATTAAT	211678
Qy	840	GGGATATTAAGGAAAGCAACCCCTTTGAGAGCAAAAGAAATGGGTTTCATATGGTGGC	899
Dh	211675	GGTGTATATGGTATCTGGAAACCATCTTAAGTATGAGAGATTAATTTTACAGTACGCT	211616
Qy	900	AAATCTATTATTGATGAATTTTGGAAAGAGT-TTACATACATCACTTAAACCCGAC	958
Dh	211615	AAAAATATTATTGAGAAATTTTCTGTGATATCTTTAAACACATATTGATTGTTGAGA	211566
Qy	959	TGCTAATGGAGGTACACATATGTGAAA-----ATGATTAATGGTCAGGGCTATATA	1010
Dh	211555	ACTCTTAATAATTCACAAAATTCGACATACATTTATCCAAAATATGATGGTCAAGC	211498
Qy	1011	ACTCAGAAATCAGGATATACATCAGAAATTTAAATTTACCTTAGCAAAATATGAGTTACT	1070
Dh	211495	ACATATTAACATCTGAGAGCTGGTAAATATTATTAATCAACTTGCATATACATTTACT	211436
Qy	1071	TTGAAAGAGAGATTAAGTTCATTAATCTTGATATGACGACCTAAATATTATCTCA	1130
Dh	211435	ATGAAAGGACAGAGCCGTGCTATCT-----AATGGTGGTGTAAACAATTTCTCC	211382
Qy	1131	CGTTAAACATGAGAAACGTATZTTTTATGGATCCAAAACAGAGTCATTAATCTTC	1130
Dh	211381	GCATTAATGTCACGGGAAAACCTTTATTTATGATTAACGAAATGGGTCGTAACTTTT	211322
Qy	1191	GCATCTGACATTAACAAAGGGGCGGTGTCTTTATTTTGAAGGTAACTTTACGTATCT	1250
Dh	211321	AGTATTAATATGATCAAGGTGCGAGCGCTTGATTTTGAAGGAGATTTTAAATTG	211262
Qy	1251	CCAAATTTTAACAAATCTGGGCAAGGACGTGAGCAACATGTAATGCAACCGTT	1310
Dh	211261	CTTACCCATTAAGAAAGTGTAAAGGGGCGGGACATCATGATGATGAAGACCGTA	211202
Qy	1311	ACTTGAAGATTAATGCGTGGACATCATGATGACCTTTCAAATAATGGTAAAGGACATTTG	1370
Dh	211201	ATTGGAAAGTCAATGSGGTGAAGAAATATGCTTCTTAATATGGTATTAAGAAAGTTA	211142
Qy	1371	CACGTTTAAGCCAAAGGGGAAATTAAGGTTGCATACGCTTAGGCGATGCTAAAGTAT	1430
Dh	211141	CATATTCAGAAAGGTGTAAATCTTAGGCTCAATATGCGTTGGCATGTGTAAAGTATC	211082
Qy	1431	TTGAGAGACAGGACGACATCATAGGCAACAAACGCTTTATGTAATTGGCTGGTT	1490
Dh	211081	TTGAAACACAGAGGACATCATATTAATTAACAAAGCTTTATGTAATTCGATTTGGTT	211022
Qy	1491	AGCGGACAGGAGCTGTTCATTTAAACATGATTAACAATTTGATACCAATTAATTTAT	1550
Dh	211021	AGCGGACAGGAGCTGTTCATTTAAACATGATTAACAATTTGATACCGATTAATTTAT	210962
Qy	1551	TTGGCTTCCTGGTGGTGGTCTTAATCTTAAAGGCAATCATTAACCTTTAAAGTATC	1610
Dh	210961	TTGGCTTCCTGGTGGTGGTCTTAAATCTTAAAGGCAATCATTAACCTTTAAAGTATC	210902
Qy	1611	CAAAATACGACAGAGGGGGCAATGATTTGTGAACATTAATACATCAAGCCGCTATGTC	1670
Dh	210901	CAAAATACGACAGAGGGGGCAATGATTTGTGAACATTAATACATCAAGTCCCTATATGTC	210842
Qy	1671	ACTATTAATCTGGGAAACGAAGCATTTGTCTCATATGCAATTAATATTAATTAATTAAT	1730
Dh	210841	ACTATTAATCTGGGAAACGAAGCATTTGTCTCATATGCAATTAATATTAATTAATTAAT	210782
Qy	1731	TACAGAAAAGAAATTCGTCACAAAGTGTGGTTGGGAAACAGATAAATTAATTAACAT	1790
Dh	210781	TACAGAAAAGAAATTCGTCACAAAGTGTGGTTGGGAAACAGATTAATTAATTAACAT	210722
Qy	1791	GGGCGATTAACCTTATTATTAACCAACACAGAAATGCTACTTGTCTACTTCAAGT	1850
Dh	210721	GGACGATTAACCTTATTATTAAACCAACACAGAAATGCTACTTGTCTACTTCAAGT	210662


```

Db 208528 TTACCCCGACATATATACGCTTAAGCTTATTTCTGTAATATATATATATATAT 208469
Qy 4053 TCACAGCTACGACACACACAGGTAATATCTACGCTTTCACACACACATTTGACGT 4112
Db 208468 TCACAGCTACGACACACACAGGTAATATCTACGCTTTCACACACACATTTGACGT 208409
Qy 4113 TATGGCAAAAAGAGTGGATTAAGCAGAAATTTTACATTTCCAAATTTCCGCTTT 4172
Db 208408 TATGGCAAAAAGAGTGGATTAAGCAGAAATTTTACATTTCCAAATTTCCGCTTT 208349
Qy 4173 ATCTCAAAATCTCAGGTTTACAGTTCGCGCAACACCAAAATGCGCGTGAATTTGCGC 4232
Db 208348 ATCTCAAAATCTCAGGTTTACAGTTCGCGCAACACCAAAATGCGCGTGAATTTGCGC 208289
Qy 4233 TATGGTGTAAATATCAATATTTATTCGTTTATTCGATTAACAGAGTGGCTCAATC 4292
Db 208288 TATGGTGTAAATATCAATATTTATTCGTTTATTCGATTAACAGAGTGGCTCAATC 208229
Qy 4293 AGATCCACCTTTTATTCGATTAAT 4319
Db 208228 AATCCACCTTTTATTCGATTAAT 208202

```

```

RESULT 3
PCT-US04-06615-5
; Sequence 5, Application PC/TUS0406615
; GENERAL INFORMATION:
; APPLICANT: New England Medical Center Hospitals, Inc.
; APPLICANT: Plant, Andrew G
; APPLICANT: Qiu, Jiazhou
; TITLE OF INVENTION: Treatment of IgA1 Deposition Diseases
; FILE REFERENCE: 28154/2268
; CURRENT APPLICATION NUMBER: PCT/US04/06615
; PRIOR FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 60/453055
; PRIOR FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 5085
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
PCT-US04-06615-5

```

```

Query Match 4.7%; Score 202; DB 1; Length 5085;
Best Local Similarity 60.9%; Pred. No. 6.1e-40;
Matches 347; Conservative 0; Mismatches 220; Indels 3; Gaps 2;

```

```

Qy 1176 GATCATTAATCTTCGATCTGACATTACCAAGGCGGCTGCTTATTTGAGGCT 1235
Db 1183 GGAACGCTTACCTTAATATATATGATCAAGGCGGCGCTTGTCTTGAAGA 1242
Qy 1236 AATTTACGATCTTCCAAATTC--TAAACCAACTTGGCAGAGTGGCAATATGTA 1292
Db 1243 GATTAATGATTAAGGCACTTGTATGATGACACTGGAAGAGTGGCTTCTT 1302
Qy 1293 AGTAAATATGACCGCTTACTTGAAGTAAATGCGTGAACATGATGCACTTTCAA 1352
Db 1303 GGTATGAAAAACGTAACGTGAAAGTACATTAACCGCAATGTATGCTTACTTAA 1362
Qy 1353 ATTGTTAAAGAAATGCGTTCACGCCAAAGGCGAAATTAAGGTTGATCGCGTA 1412
Db 1363 ATCGCAAGAAACATTAATTTAGAGGAAAGGAGAAATTAAGGTTGCTAAAGTG 1422
Qy 1413 GCGATGTTAAAGTCAATTTTGAGCAGCGGCAAGCATCAAGGCAACAAAGACCTTT 1472
Db 1423 GCGATGTTAAAGTCAATTTTGAGCAGCGGCAAGCATCAATTAAGTAAAGCTTT 1482
Qy 1473 AGTAAATGCGCTTGGTTAGCGGCAAGGAGCTGTTCAATTAACGATTAACAAATTT 1532
Db 1483 TCACAGTAGTATAGTAAAGTGTGCTCAACGTTGTAATTAATGATTAAGCAAGTA 1542

```

```

Qy 1533 GATACGATTAATTTTATTTGCGCTTGGTGGTGGCTTATGATCTTAACGGGATTC 1592
Db 1543 GATCAAAATTCATTTATTTGCTTTAGAGGTGGTGGATTAATCCAAATGGCAATAT 1602
Qy 1593 TTAACCTTTAAAGTATCCAAATTAAGGAGCGAGGCGAAATGATTTGACCATTAACA 1652
Db 1603 CTCACCTTTGAACATTCGTAATTTGATGATGATGATGATGATGATGATGATGATG 1662
Qy 1653 ACTGAGCGGTAAATGATCACTTACTTGGGAGCAAGAGCTGCTTACATTAAGAAAT 1712
Db 1663 AGCAAACTCTTACTGTAACATTTACTGCGGAAAGCTTAATTAAGATCAATTAAT 1722
Qy 1713 AATATTAATTAACCTGATTAACAGAAAGA 1742
Db 1723 ACTGATTAATTAAGAGCAACCATGA 1752

```

```

RESULT 4
PCT-US04-07001-563/c
; Sequence 563, Application PC/TUS0407001
; GENERAL INFORMATION:
; APPLICANT: BAKARSTZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: PCT/US04/07001
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 563
; LENGTH: 14710
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8250)..(8250)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8293)..(8293)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8350)..(8350)
; OTHER INFORMATION: n = a, c, g, or t
PCT-US04-07001-563

```

```

Query Match 4.1%; Score 177.6; DB 1; Length 14710;
Best Local Similarity 59.3%; Pred. No. 1.1e-33;
Matches 341; Conservative 0; Mismatches 225; Indels 9; Gaps 2;

```

```

Qy 1154 ATATTTATGATCAAAAACAGATCAATATCTTGCATCTGACATTAACCAAGGCG 1213
Db 8261 AATATGATGATTAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8202
Qy 1214 GGGTGGCTTATTTTATGAGGTAATTTTACAGTATCCAAATTC--TAAACCAACTTG 1270
Db 8201 AGGCGATTTGCTTTTGAAGTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8142
Qy 1271 GCAAGAGCTGGCATACATATTAAGTAAATTAAGCAACGCTTACTGGAAGTAAATG 1330
Db 8141 GAAAGGAGGCTGGTATATCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8082
Qy 1331 GGAACATATGACCTTTCTTAATTTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1390
Db 8081 TCAATTTGATGCTTATTAATTTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8022
Qy 1391 AATTAAGGTTGATTCAGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1450
Db 8021 TATTAAGGTTGATTCAGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7962
Qy 1451 TCAGGCAACAAAGACCTTTAGTAAATTTGGCTTGTAGGCGAGAGGAGGAGGAGGAG 1510

```

```

Db      7961 AGGAC-----AACAGCTTTGCTCTGAGGAATGTAAGTGGTGGCTCAACTGTTGT 7908
Qy      1511 ATTAAGATGATTAACAACTTTGATACCGATTAATTTTATTTGCGCTTCGTGGTGG 1570
Db      7907 ACTTAAGATGATTAACAACTTTGATACCGATTAATTTTATTTGCGCTTCGTGGTGG 7848
Qy      1571 CTGATCTTAACGGGCAATTCATTAACCTTAAGATTAAGTATCCAAATATACGAGAGGAGG 1630
Db      7847 ATTAGATGCCAATGCGAATATCTCACTTTTGAACATATCCGTAATTTTGAATGATGCTGC 7788
Qy      1631 AATGATTTGAACCAATATATACACTCAAGCCGTAATGTCATTTATCTGGAGAACGAAG 1690
Db      7787 AAGACTAGTATATCAATATATGACTATGCTCAATATATACGATTAATCTGAGGACGAGATT 7728
Qy      1691 CATGCTTACCTTAATGGAATATATATATTAAC 1725
Db      7727 AATTACAAATCCAAAGTCAGTCACTATATATACAC 7693

```

RESULT 5

PCT-US04-07001-762
Sequence 762, Application PC/TUS0407001

```

; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: PCT/US04/07001
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 762
; LENGTH: 5385
; TYPE: DNA
; ORGANISM: H. influenzae
; PCT-US04-07001-762

```

Query Match

Best Local Similarity 59.3%; Pred. No. 1,1e-33; Length 5385;
Matches 341; Conservative 0; Mismatches 225; Indels 9; Gaps 2;

```

Qy      1154 ATATTATGATCAAAACAAAGATCAATTAATCTTGCATCTGACATTAACCAAGGGGC 1213
Db      1248 AATGTTAGCTTGAAGAGAGTGAACCTTACTTAATTAATGATCAAGTGC 1307
Qy      1214 GGGTGTCTTATTTTGAAGGTATTTTACAGTATCTCCAAATTC---TAACCAACTTG 1270
Db      1308 AGCGGATGTGTTCTTGAAGGTGATATGAGTAAAGGCACCTTCAGAAATATGACCTTG 1367
Qy      1271 GCAGGAGCTGCGATACATGTAAGTGAATAGCCGCTACTTCTGAAAGTAAATGCGGT 1330
Db      1368 GAAAGGGGCTGTAATCTGTGCTGAAGAAAAACATTAATGTGAAGAGTGCATTAACCC 1427
Qy      1331 GGAACATGATCACTTTCTAAATTTGGTAAAGAACATTCAGCTTCAGCCAAAGGGGA 1390
Db      1428 TCATTTGATCGTTAGCTAAATTTGGTAAAGAGTAAATTTGTTGAGAGAGAGA 1487
Qy      1391 AATTAAGTTGATCAAGCGTGAAGTGAATGTCATTTTGAAGCAAGCAAGCAAGCA 1450
Db      1488 TAAATAAAGTTGCTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1547
Qy      1451 TCAGGCAACAAAGCCTTTAGTGAATTTGGCTGTTAGAGGAGGAGGAGCTGTCA 1510
Db      1548 AGGAC-----AACAGCTTTTGTCTCTGAGGAATGTAAGTGGTGGCTCAACTGTTGT 1601
Qy      1511 ATTAAGATGATTAACAACTTTGATACCGATTAATTTTATTTGCGCTTCGTGGTGG 1570
Db      1602 ACTTAAGATGATTAACAACTTTGATACCGATTAATTTTATTTGCGCTTCGTGGTGG 1661
Qy      1571 CTGATCTTAACGGGCAATTCATTAACCTTAAGATTAAGTATCCAAATATACGAGAGGAGG 1630

```

```

Db      1662 ATTAGATGCCAATGCGAATATCTCACTTTGAACATATCCGTAATTTGATGATGCTGC 1721
Qy      1631 AATGATTTGAACCAATATATACACTCAAGCCGTAATGTCATTTACTGGAAACGAAG 1690
Db      1722 AAGACTAGTATATCAATATATGACTAATGCTCAATATATACGATTAATCTGAGGACGAGATT 1781
Qy      1691 CATGCTTACCTTAATGGAATATATATTAAC 1725
Db      1782 AATTACAAATCCAAAGTCAGTCACTATATATACAC 1816

```

RESULT 6

PCT-US04-07001-683/c
Sequence 683, Application PC/TUS0407001

```

; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: PCT/US04/07001
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 683
; LENGTH: 191996
; TYPE: DNA
; ORGANISM: H. influenzae
; PCT-US04-07001-683

```

Query Match

Best Local Similarity 59.3%; Pred. No. 4.3e-33; Length 191996;
Matches 341; Conservative 0; Mismatches 225; Indels 9; Gaps 2;

```

Qy      1154 ATATTATGATCAAAACAAAGATCAATTAATCTTGCATCTGACATTAACCAAGGGGC 1213
Db      170400 AATGTTAGCTTGAAGAGAGTGAACCTTACTTAATTAATGATCAAGTGC 170341
Qy      1214 GGGTGTCTTATTTTGAAGGTATTTTACAGTATCTCCAAATTC---TAACCAACTTG 1270
Db      170340 AGCGGATGTGTTCTTGAAGGTGATATGAGTAAAGGCACCTTCAGAAATATGACCTTG 170281
Qy      1271 GCAGGAGCTGCGATACATGTAAGTGAATAGCCGCTACTTCTGAAAGTAAATGCGGT 1330
Db      170280 GAAAGGGGCTGTAATCTGTGCTGAAGAAAAACATTAATGTGAAGAGTGCATTAACCC 170221
Qy      1331 GGAACATGATCACTTTCTAAATTTGGTAAAGAACATTCAGCTTCAGCCAAAGGGGA 1390
Db      170220 TCATTTGATCGTTAGCTAAATTTGGTAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 170161
Qy      1391 AATTAAGTTGATCAAGCGTGAAGTGAATGTCATTTTGAAGCAAGCAAGCAAGCA 1450
Db      170160 TAAATAAAGTTGCTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 170101
Qy      1451 TCAGGCAACAAAGCCTTTAGTGAATTTGGCTGTTAGAGGAGGAGGAGCTGTCA 1510
Db      170100 AGGAC-----AACAGCTTTTGTCTCTGAGGAATGTAAGTGGTGGCTCAACTGTTGT 170047
Qy      1511 ATTAAGATGATTAACAACTTTGATACCGATTAATTTTATTTGCGCTTCGTGGTGG 1570
Db      170046 ACTTAAGATGATTAACAACTTTGATACCGATTAATTTTATTTGCGCTTCGTGGTGG 169987
Qy      1571 CTGATCTTAACGGGCAATTCATTAACCTTAAGATTAAGTATCCAAATATACGAGAGGAGG 1630
Db      169986 ATTAGATGCCAATGCGAATATCTCACTTTTGAACATATCCGTAATTTGATGATGCTGC 169927
Qy      1631 AATGATTTGAACCAATATATACACTCAAGCCGTAATGTCATTTACTGGAGAACGAAG 1690
Db      169926 AAGACTAGTATATCAATATATGACTAATGCTCAATATATACGATTAATCTGAGGACGAGATT 169867
Qy      1691 CATGCTTACCTTAATGGAATATATTAAC 1725
Db      169866 AATTACAAATCCAAAGTCAGTCACTATATATACAC 169832

```

RESULT 7
US-10-215-432A-29

Sequence 29, Application US/10215432A

GENERAL INFORMATION:

APPLICANT: KMEC, ERIC B.

APPLICANT: PAREKH-OLMEDO, HETAL

TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE PREVENTION AND TREATMENT OF

TITLE OF INVENTION: HUNTINGTON'S DISEASE

FILE REFERENCE: NaPro-10 US

CURRENT APPLICATION NUMBER: US/10/215,432A

CURRENT FILING DATE: 2002-08-07

PRIOR APPLICATION NUMBER: US 60/310,757

PRIOR FILING DATE: 2001-08-07

PRIOR APPLICATION NUMBER: US 60/310,889

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/310,770

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/337,219

PRIOR FILING DATE: 2001-12-04

NUMBER OF SEQ ID NOS: 63

SOFTWARE: PatentIn ver. 3.2

SEQ ID NO 29

LENGTH: 561

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: Repeat region

LOCATION: (51) (411)

OTHER INFORMATION: this region may encompass 20-120 car repeats according to

OTHER INFORMATION: the specification as filed; see specification for detailed

US-10-215-432A-29

Query Match 1.1%; Score 48.6; DB 7; Length 561;

Best Local Similarity 34.0%; Pred. No. 0.047;

Matches 101; Conservative 63; Mismatches 132; Indels 1; Gaps 1;

QY 2953 TGAAGATGATGCGGATTCGCTTGGCAACCAATATAAGAGAGAGATTCACATG 3012

DB 20 TATGAGCGCTTGAGTCCCTCAAGTCTTCCARCARCARCARCARCARCARCAR 79

QY 3013 ATTGTAG 3072

DB 80 A-RCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCAR 138

QY 3073 CTAAAGCAAAAG 3132

DB 139 CACARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCAR 198

QY 3133 CTGATACCTGCTGATCAAGCTGTAAGCATTAGAGAGAGAGAGAGAGAGAGAG 3192

DB 199 CACARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCAR 258

QY 3193 CTGCTGAAG 3249

DB 259 CACARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCAR 315

RESULT 8

US-10-859-198-1510

Sequence 1510, Application US/10859198

GENERAL INFORMATION:

APPLICANT: Wyeth

APPLICANT: Mounts, William

APPLICANT: Murphy, Ellen

APPLICANT: Whitley, Maryann

TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a Non-Viral

TITLE OF INVENTION: Species

FILE REFERENCE: 031896-014000 (AM101085)

CURRENT APPLICATION NUMBER: US/10/859,198

CURRENT FILING DATE: 2004-06-03

PRIOR APPLICATION NUMBER: US 60/475,871

PRIOR FILING DATE: 2003-06-05

NUMBER OF SEQ ID NOS: 282011

SOFTWARE: PatentIn version 3.2

SEQ ID NO 1510

LENGTH: 4647

TYPE: DNA

ORGANISM: Staphylococcus aureus

US-10-859-198-1510

Query Match 1.1%; Score 48.2; DB 7; Length 4647;

Best Local Similarity 44.3%; Pred. No. 0.13;

Matches 197; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 2988 ATAAAG 3047

DB 319 ACAAAG 378

QY 3048 GAG 3107

DB 379 GAG 438

QY 3108 AG 3167

DB 439 GAG 498

QY 3168 TTAAAG 3227

DB 439 GAG 558

QY 3228 AAG 3287

DB 559 GAG 618

QY 3288 TTAAAG 3347

DB 619 ACAAAG 678

QY 3348 CAG 3407

DB 679 GCTGAG 738

QY 3408 GCGTAAAG 3467

DB 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

QY 739 GTAG 763

```
QY 2988 ATAAAGAGCAGGATTCGATGATTTAGTAGAGCAGACAGACGACATTA 3047
    |||||
DB 319 ACAACTGACACAGCATCAACAGAGAGAGAGAGATCACTGACAGACACACAGAA 378
QY 3048 GAGCCCAACAGAGTGAACCGATCTGTAACACACAGAGTGAAGCCAAAGTGCGTCA 3107
    |||||
DB 379 GAGCGCCCAAGCTGAGAGACAGACAGAGTGAACAGAGAGAGCCCAAAAGCTGAA 438
QY 3108 AGAAGAGCAGCAGAGCAGCGTTTCTGATACCGCTGATGAAGCCTGTTAAAGCA 3167
    |||||
DB 439 GAAACAGACAAAGCAGACAGAGAGAGAGCCAAAGCTGAGAGAGACAGACAGAA 498
QY 3168 TTAAAGCCCAACAGAGTGAACCTGATCTGTAACACACAGAGTGAAGCCAAAGCA 3227
    |||||
DB 499 GAGAGCGCCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
QY 3228 AAAGTGGCGTCAAAAGAGCAGTGTCTTCTGATCCCTGCTTGAACAGCCTGTTGCA 3287
    |||||
DB 559 GAAACAGACAAAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618
QY 3288 TTAAAGCGCGCACTTGAAGTATGATGCCCAAGCAGCATGCAAAAGATGCTAGCT 3347
    |||||
DB 619 ACAGAGAGAGCGCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678
QY 3348 CAAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
    |||||
DB 679 GCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
QY 3408 GCGTTATCAGATTAATCTGCAACAG 3432
    |||||
DB 739 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 763
```

```
RESULT 10
US-10-859-198-1512
; Sequence 1512, Application US/10859198
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; APPLICANT: Murphy, Ellen
; TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a Non-Viral
; FILE REFERENCE: 031896-014000 (AM101085)
; CURRENT APPLICATION NUMBER: US/10/859,198
; CURRENT FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US 60/475,871
; PRIOR FILING DATE: 2003-06-05
; NUMBER OF SEQ ID NOS: 282011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1512
; LENGTH: 5097
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-859-198-1512
```

```
Query Match 1.1%; Score 48.2; DB 7; Length 5097;
Best Local Similarity 44.3%; Pred. No. 0.14;
Matches 197; Conservative 0; Mismatches 248; Indels 0; Gaps 0;
```

```
QY 2988 ATAAAGAGCAGGATTCGATGATTTAGTAGAGCAGACAGACGACATTA 3047
    |||||
DB 319 ACAACTGACACAGCATCAACAGAGAGAGAGAGATCACTGACAGACACACAGAA 378
QY 3048 GAGCCCAACAGAGTGAACCGATCTGTAACACACAGAGTGAAGCCAAAGTGCGTCA 3107
    |||||
DB 379 GAGCGCCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 438
QY 3108 AGAAGAGCAGCAGAGCAGCGTTTCTGATACCGCTGATGAAGCCTGTTAAAGCA 3167
    |||||
DB 439 GAAACAGACAAAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 498
```

```
QY 3168 TTAAAGCCCAACAGAGTGAACCTGATCTGTAACACACAGAGTGAAGCCAAAGCA 3227
    |||||
DB 499 GAGAGCGCCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
QY 3228 AAAGTGGCGTCAAAAGAGCAGTGTCTTCTGATCCCTGCTTGAACAGCCTGTTGCA 3287
    |||||
DB 559 GAAACAGACAAAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618
QY 3288 TTAAAGCGCGCACTTGAAGTATGATGCCCAAGCAGCATGCAAAAGATGCTAGCT 3347
    |||||
DB 619 ACAGAGAGAGCGCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678
QY 3348 CAAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
    |||||
DB 679 GCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
QY 3408 GCGTTATCAGATTAATCTGCAACAG 3432
    |||||
DB 739 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 763
```

```
RESULT 11
US-10-859-198-9364
; Sequence 9364, Application US/10859198
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; APPLICANT: Murphy, Ellen
; TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a Non-Viral
; FILE REFERENCE: 031896-014000 (AM101085)
; CURRENT APPLICATION NUMBER: US/10/859,198
; CURRENT FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US 60/475,871
; PRIOR FILING DATE: 2003-06-05
; NUMBER OF SEQ ID NOS: 282011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9364
; LENGTH: 5097
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-859-198-9364
```

```
Query Match 1.1%; Score 48.2; DB 7; Length 5097;
Best Local Similarity 44.3%; Pred. No. 0.14;
Matches 197; Conservative 0; Mismatches 248; Indels 0; Gaps 0;
```

```
QY 2988 ATAAAGAGCAGGATTCGATGATTTAGTAGAGCAGACAGACGACATTA 3047
    |||||
DB 319 ACAACTGACACAGCATCAACAGAGAGAGAGAGATCACTGACAGACACACAGAA 378
QY 3048 GAGCCCAACAGAGTGAACCGATCTGTAACACACAGAGTGAAGCCAAAGTGCGTCA 3107
    |||||
DB 379 GAGCGCCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 438
QY 3108 AGAAGAGCAGCAGAGCAGCGTTTCTGATACCGCTGATGAAGCCTGTTAAAGCA 3167
    |||||
DB 439 GAAACAGACAAAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 498
QY 3168 TTAAAGCCCAACAGAGTGAACCTGATCTGTAACACACAGAGTGAAGCCAAAGCA 3227
    |||||
DB 499 GAGAGCGCCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
QY 3228 AAAGTGGCGTCAAAAGAGCAGTGTCTTCTGATCCCTGCTTGAACAGCCTGTTGCA 3287
    |||||
DB 559 GAAACAGACAAAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618
QY 3288 TTAAAGCGCGCACTTGAAGTATGATGCCCAAGCAGCATGCAAAAGATGCTAGCT 3347
    |||||
DB 619 ACAGAGAGAGCGCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678
QY 3348 CAAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
    |||||
```

TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a

Query Match	1.0%; Score 44.4; DB 1; Length 1543;

Best Local Similarity 50.0%; Pred. No. 0.77;

Matches 111; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

```
QY 3800 TCATCGAAGAGCATTAATTATGCGGTAATGCAAGTATCATGTTCCGTTAGGGCAATT 3859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1068 TCATGGAAGAAAAATATATATAGATTGAGTGAAGATTAAGGGGTCTAGAAATTTT 1127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3860 GGGCAATTCAGCTTATTTTGAGTTAATCGCTATTTTATTAAGTGAAGTAATTCATC 3919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1128 CTGATTAATCTGCTGCAATTTCTGATTAATTTCTGATTTCTGTAAGTATTAATTTT 1187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3920 TGAGGAGTGAAGTGAAGAAAGCGCTGACCTTGATTTAATGCTATTAATGCGCATCG 3979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1188 ATAAAGATTAAAGAAAGCTCTATAGAACTATAGTAGTAGCTGACGAGAGAAATTA 1247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3980 AGTTGATTATACATTACTCGACGATATATACGCGTTAA 4021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1248 ACTTGAAATTAATTAATTAATTAATTAATTAATTAATTA 1289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

PCT-US03-41761-38500
; Sequence 38500, Application PC/TUS0341761
; GENERAL INFORMATION:
; APPLICANT: NMT GENOMICS, INC.
; APPLICANT: DENISER, Sue K.
; APPLICANT: CHARTERIS, Paul
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND SYSTEMS FOR INFERRING BOVINE BREED
; FILE REFERENCE: NM1150M0
; CURRENT APPLICATION NUMBER: PCT/US03/41761
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent version 3.1
; SEQ ID NO 38500
; LENGTH: 1543
; TYPE: DNA
; ORGANISM: Bovine 19866801025522
PCT-US03-41761-38500

Query Match

1.0%; Score 44.4; DB 1; Length 1543;
Best Local Similarity 50.0%; Pred. No. 0.77;

Matches 111; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

```
QY 3800 TCATCGAAGAGCATTAATTATGCGGTAATGCAAGTATCATGTTCCGTTAGGGCAATT 3859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1068 TCATGGAAGAAAAATATATATAGATTGAGTGAAGATTAAGGGGTCTAGAAATTTT 1127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3860 GGGCAATTCAGCTTATTTTGAGTTAATCGCTATTTTATTAAGTGAAGTAATTCATC 3919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1128 CTGATTAATCTGCTGCAATTTCTGATTAATTTCTGATTTCTGTAAGTATTAATTTT 1187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3920 TGAGGAGTGAAGTGAAGAAAGCGCTGACCTTGATTTAATGCTATTAATGCGCATCG 3979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1188 ATAAAGATTAAAGAAAGCTCTATAGAACTATAGTAGTAGCTGACGAGAGAAATTA 1247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3980 AGTTGATTATACATTACTCGACGATATATACGCGTTAA 4021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1248 ACTTGAAATTAATTAATTAATTAATTAATTAATTAATTA 1289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: June 24, 2004, 19:55:42
Job time : 670 secs

1 TCAATAGTCGTTTAACTA

```

Db      1  TCAATAGTCGTTTACAGATTTTATATACAAAATTTCTTAATTAATAAATTAACATTA 60
Qy      61  TGAATAAAACGTAATTCGCTTAATTTTAAACCGCTGCAATTCATTAAGGATATGAT 120
Db      61  TGAATAAAACGTAATTCGCTTAATTTTAAACCGCTGCAATTCATTAAGGATATGAT 120
Qy      121  CGCAAGCGTGGGCTGGTCACTTATTTTGGAGTTGATTAACCAATTAATTCGATATTTG 180
Db      121  CGCAAGCGTGGGCTGGTCACTTATTTTGGAGTTGATTAACCAATTAATTCGATATTTG 180
Qy      181  CGGAGATTAAGGGAAGTTCAAGATGGGGCTCAAAATATTAAGTTTAAACAAACAG 240
Db      181  CGGAGATTAAGGGAAGTTCAAGATGGGGCTCAAAATATTAAGTTTAAACAAACAG 240
Qy      241  GGCATTTGTTGGCAATCAATGACAAACCCCAATGATGATTTTTCGTAAGTAC 300
Db      241  GGCATTTGTTGGCAATCAATGACAAACCCCAATGATGATTTTTCGTAAGTAC 300
Qy      301  GTAACGGCGTGGACGCTTGG-TGAAATCAATATATGTAGCGTGGCAATTAACGATG 360
Db      301  GTAACGGCGTGGACGCTTGGTGAATTAATATATGTAGCGTGGCAATTAACGATG 360
Qy      361  GATATACGATTTGATTTTGGTGAAGGGAACACCCGATCAACATCGTTTACTT 420
Db      361  GATATACGATTTGATTTTGGTGAAGGGAACACCCGATCAACATCGTTTACTT 420
Qy      421  ATAAGAATTGAAAGAAATCACTACAAAAGATTAATTAATCCTCTTAAGGACAT 480
Db      421  ATAAGAATTGAAAGAAATCACTACAAAAGATTAATTAATCCTCTTAAGGACAT 480
Qy      481  ACCATAATCAACGATTAATTAATTCGTTACAGAAACGGCTCCAAATGATATGACTTGA 540
Db      481  ACCATAATCAACGATTAATTAATTCGTTACAGAAACGGCTCCAAATGATATGACTTGA 540
Qy      541  ATATGATGCGACGACTTATTCGATTAAGAAACAAATATTCGAAACGTTTGGTATCGCT 600
Db      541  ATATGATGCGACGACTTATTCGATTAAGAAACAAATATTCGAAACGTTTGGTATCGCT 600
Qy      601  CTGGACGGGAGTTTGGGGAATGATCAACAAAGGAGCAACAGTTGCGGTGCAATAC 660
Db      601  CTGGACGGGAGTTTGGGGAATGATCAACAAAGGAGCAACAGTTGCGGTGCAATAC 660
Qy      661  ATATCTGACAGCTGGCAATACACACATCAGCGTGGAGCAAGTAATGATTTGCTAT 720
Db      661  ATATCTGACAGCTGGCAATACACACATCAGCGTGGAGCAAGTAATGATTTGCTAT 720
Qy      721  TGGAGGCGATTTGCTAAAGCGGAGATATGCTCATTCGATTCGAGTTCGACCTCAAG 780
Db      721  TGGAGGCGATTTGCTAAAGCGGAGATATGCTCATTCGATTCGAGTTCGACCTCAAG 780
Qy      781  GGGACAGTGTCTCGAGATTTATTAATGATGCTGAAGAAACAAAATGTTAATATG 840
Db      781  GGGACAGTGTCTCGAGATTTATTAATGATGCTGAAGAAACAAAATGTTAATATG 840
Qy      841  GGAATATTACGGGAAGGCAACCTTTTGAAGGCAAGAAATGAGTTTCAATGCTTGC 900
Db      841  GGAATATTACGGGAAGGCAACCTTTTGAAGGCAAGAAATGAGTTTCAATGCTTGC 900
Qy      901  AATCTATTTTGAAGAAATTTTGAAGAGATTTACATACATCACTTTCACCCGAGCTG 960
Db      901  AATCTATTTTGAAGAAATTTTGAAGAGATTTACATACATCACTTTCACCCGAGCTG 960
Qy      961  GTAATGAGTGTACAAATAGTGAATGATATGTCAGGGCTATTAATCAAGAT 1020
Db      961  GTAATGAGTGTACAAATAGTGAATGATATGTCAGGGCTATTAATCAAGAT 1020
Qy      1021  CAGAAATCCATCAAGAAATTAATTAATGCTTGAAGCAATATGATTTACCTTTGAAGGA 1080
Db      1021  CAGAAATCCATCAAGAAATTAATTAATGCTTGAAGCAATATGATTTACCTTTGAAGGA 1080
Qy      1081  AGGATTAAGTTCAATCTGATATGAGATGAGACCTAATATTAATTCCTCGCTTTAAACA 1140
Db      1081  AGGATTAAGTTCAATCTGATATGAGATGAGACCTAATATTAATTCCTCGCTTTAAACA 1140

Db      1281  AGGATTAAGTTCAATCTGATATGAGATGAGACCTAATATTAATTCCTCGCTTTAAACA 1140
Qy      1141  ATGGAATAAGCCATTAATTTATGATGATCAAAAACAAAGATCATTAATCTTCGATCTGACA 1200
Db      1141  ATGGAATAAGCCATTAATTTATGATGATCAAAAACAAAGATCATTAATCTTCGATCTGACA 1200
Qy      1201  TTAACCAAGGGGCGGGTGTCTTATTTTGAAGGTAATTTTACAGATCTCCAAATTTCA 1260
Db      1201  TTAACCAAGGGGCGGGTGTCTTATTTTGAAGGTAATTTTACAGATCTCCAAATTTCA 1260
Qy      1261  ACCAAACTTGGCAAGGCTGGCATCATGTAGTGAATAACACCGTTACTTTGGAAG 1320
Db      1261  ACCAAACTTGGCAAGGCTGGCATCATGTAGTGAATAACACCGTTACTTTGGAAG 1320
Qy      1321  TAAATGGCTGGAAACATGATCGACTTTGAAATTTGTTAAAGGAACTTTCAGCTTAA 1380
Db      1321  TAAATGGCTGGAAACATGATCGACTTTGAAATTTGTTAAAGGAACTTTCAGCTTAA 1380
Qy      1381  CCAAAAGGGGAAATTAAGTTCGATCAGGTAAGGCGATGGTAAGTCACTTTTGGACAGC 1440
Db      1381  CCAAAAGGGGAAATTAAGTTCGATCAGGTAAGGCGATGGTAAGTCACTTTTGGACAGC 1440
Qy      1441  AGGACAGCATCAAGGCAACAAACAGCTTTAGTAATTTGCTTGGTGGGAGAG 1500
Db      1441  AGGACAGCATCAAGGCAACAAACAGCTTTAGTAATTTGCTTGGTGGGAGAG 1500
Qy      1501  GGACTGTCATTAAGGATTAACGATTAACCTTGAATTTGATTCGCTTTC 1560
Db      1501  GGACTGTCATTAAGGATTAACGATTAACCTTGAATTTGATTCGCTTTC 1560
Qy      1561  GTGGTGGTGGCTGATCTTAACGGGCACTTCAATTAACCTTTAAACGTAATCCAAATACG 1620
Db      1561  GTGGTGGTGGCTGATCTTAACGGGCACTTCAATTAACCTTTAAACGTAATCCAAATACG 1620
Qy      1621  ACCAGGGGGCAATGATTTGTAACCAATTAACACTCAAGCCCTAATGCTCATTAATG 1680
Db      1621  ACCAGGGGGCAATGATTTGTAACCAATTAACACTCAAGCCCTAATGCTCATTAATG 1680
Qy      1681  GGAAGGAAGCAATGTTCTACCTTAATGGAATTAATTAATTAATTAATTAATTAATTA 1740
Db      1681  GGAAGGAAGCAATGTTCTACCTTAATGGAATTAATTAATTAATTAATTAATTAATTA 1740
Qy      1741  AAATTCCTCAACGATTTGCTTGGGGAACAGATTAATAATTAACAAATGGCGATTA 1800
Db      1741  AAATTCCTCAACGATTTGCTTGGGGAACAGATTAATAATTAACAAATGGCGATTA 1800
Qy      1801  ACCCTATTATTAACCAACCAACAGATGCTTGTCTTACCTTCAAGTGTGTAATTA 1860
Db      1801  ACCCTATTATTAACCAACCAACAGATGCTTGTCTTACCTTCAAGTGTGTAATTA 1860
Qy      1861  TAAAGGCGATTTACCAACCAACAAAGGTAACATTTTTCAGCGGTGACCGACCGC 1920
Db      1861  TAAAGGCGATTTACCAACCAACAAAGGTAACATTTTTCAGCGGTGACCGACCGC 1920
Qy      1921  ACGCCCTACATCAATTTAAATTAACGTTGTCAGAAATGAGATATACCAAGGCGAA 1980
Db      1921  ACGCCCTACATCAATTTAAATTAACGTTGTCAGAAATGAGATATACCAAGGCGAA 1980
Qy      1981  TTGTGGGATCAAGATTTGATTAACCGTATTAATTAAGTGAACCTTCCAAATTTAAG 2040
Db      1981  TTGTGGGATCAAGATTTGATTAACCGTATTAATTAAGTGAACCTTCCAAATTTAAG 2040
Qy      2041  GCGGAAGTGGCGGTGTTTTCGCAATGTTTCTTCAATGAGGAAATGAGACGTCAGCA 2100
Db      2041  GCGGAAGTGGCGGTGTTTTCGCAATGTTTCTTCAATGAGGAAATGAGACGTCAGCA 2100
Qy      2101  ATAAATGCAATGCAATTTTGGTGTGTCGCAATCAACAAATTAATTTGACGCGTT 2160
Db      2101  ATAAATGCAATGCAATTTTGGTGTGTCGCAATCAACAAATTAATTTGACGCGTT 2160
Qy      2161  CAGATTTGACGATTAAGCATTTGCAAAAGTGAATTAACGATTAACAAAGTTATTA 2220
Db      2161  CAGATTTGACGATTAAGCATTTGCAAAAGTGAATTAACGATTAACAAAGTTATTA 2220

```


1. *Chlorophyll a* (Chl a) is the primary photosynthetic pigment in most plants and algae. It is a green pigment that absorbs light energy in the blue and red regions of the visible spectrum. Chl a is essential for the light-dependent reactions of photosynthesis, where it converts light energy into chemical energy in the form of ATP and NADPH.

2. *Chlorophyll b* (Chl b) is an accessory pigment found in green plants and green algae. It is a yellow-green pigment that absorbs light energy in the blue and red regions of the visible spectrum. Chl b transfers the absorbed energy to Chl a, which then uses it for photosynthesis.

3. *Carotenoids* are a group of pigments that include carotenes and xanthophylls. They are responsible for the yellow, orange, and red colors seen in autumn foliage. Carotenoids absorb light energy in the blue and green regions of the visible spectrum and transfer the energy to Chl a. They also play a role in protecting the photosynthetic apparatus from damage by reactive oxygen species.

4. *Xanthophylls* are a type of carotenoid that are responsible for the yellow color seen in autumn foliage. They absorb light energy in the blue and green regions of the visible spectrum and transfer the energy to Chl a. Xanthophylls also play a role in protecting the photosynthetic apparatus from damage by reactive oxygen species.

5. *Anthocyanins* are water-soluble pigments that are responsible for the red, purple, and blue colors seen in autumn foliage. They are not involved in photosynthesis but are produced by the plant in response to environmental factors such as low temperatures and high light intensity.

6. *Flavonoids* are a group of pigments that include flavones and flavanols. They are responsible for the yellow and orange colors seen in autumn foliage. Flavonoids absorb light energy in the blue and green regions of the visible spectrum and transfer the energy to Chl a. They also play a role in protecting the photosynthetic apparatus from damage by reactive oxygen species.

7. *Anthoxanthins* are a type of flavonoid that are responsible for the yellow color seen in autumn foliage. They absorb light energy in the blue and green regions of the visible spectrum and transfer the energy to Chl a. Anthoxanthins also play a role in protecting the photosynthetic apparatus from damage by reactive oxygen species.

8. *Anthocyanins* are water-soluble pigments that are responsible for the red, purple, and blue colors seen in autumn foliage. They are not involved in photosynthesis but are produced by the plant in response to environmental factors such as low temperatures and high light intensity.

9. *Anthocyanins* are water-soluble pigments that are responsible for the red, purple, and blue colors seen in autumn foliage. They are not involved in photosynthesis but are produced by the plant in response to environmental factors such as low temperatures and high light intensity.

10. *Anthocyanins* are water-soluble pigments that are responsible for the red, purple, and blue colors seen in autumn foliage. They are not involved in photosynthesis but are produced by the plant in response to environmental factors such as low temperatures and high light intensity.

Qy	361	GATATACAGATGTTGATTTTGGTTCAGAGGAAACACCACCATCAATCGTTTACT	420
Dp	361	GATATACAGATGTTGATTTTGGTTCAGAGGAAACACCACCATCAATCGTTTACT	420
Qy	421	ATATGATGTATAAACGAATTAATCAAAAAAGATATTTACATCCTTATGAGCAGAT	480
Dp	421	ATATGATGTATAAACGAATTAATCAAAAAAGATATTTACATCCTTATGAGCAGAT	480
Qy	481	ACCATATTCACGATACATCAATATTTGGTTACAGAGGGGCTCAATATATATGACTTGA	540
Dp	481	ACCATATTCACGATACATCAATATTTGGTTACAGAGGGGCTCAATATATATGACTTGA	540
Qy	541	ATATGAAATGGCAGTACTTTCAGATAGCAAAATATTCGAAAGCGTGTGATCGCT	600
Dp	541	ATATGAAATGGCAGTACTTTCAGATAGCAAAATATTCGAAAGCGTGTGATCGCT	600
Qy	601	CTGACAGCGCACTTTTGGCGAATATATACAGCAAAAGGCGACAGTTGCGGTGATATC	660
Dp	601	CTGACAGCGCACTTTTGGCGAATATATACAGCAAAAGGCGCGACAGTTGCGGTGATATC	660
Qy	661	ATTATCTGACGCTGGCAATACACATCATCAGCGACAGGTAATGATATTCGATT	720
Dp	661	ATTATCTGACGCTGGCAATACACATCATCAGCGACAGGTAATGATATTCGATT	720
Qy	721	TGGAGGCGGATGTTTGGTAAGCGGACATATGTCTCATTCGATTCGAGGCTCAAGG	780
Dp	721	TGGAGGCGGATGTTTGGTAAGCGGACATATGTCTCATTCGATTCGAGGCTCAAGG	780
Qy	781	GGGACAGTGTCTTCGAGTATTTATATGATGCTGAAAAAGAAATGCTATATATAG	840
Dp	781	GGGACAGTGTCTTCGAGTATTTATATGATGCTGAAAAAGAAATGCTATATATAG	840
Qy	841	GGAATATTCAGGGAAGGCAACCCCTTTTGAAAGGCAAGAAATGCGTTCAATGGTTCGA	900
Dp	841	GGAATATTCAGGGAAGGCAACCCCTTTTGAAAGGCAAGAAATGCGTTCAATGGTTCGA	900
Qy	901	AATCTATTTTGAATGAAATTTTTCGAAAGGATTTACATACATCTTACACCCGAGCTG	960
Dp	901	AATCTATTTTGAATGAAATTTTTCGAAAGGATTTACATACATCTTACACCCGAGCTG	960
Qy	961	GTAATGAGGTATCAATATGATGAAATGATATAGCTCAGGGGCTATATCAAGAAAT	1020
Dp	961	GTAATGAGGTATCAATATGATGAAATGATATAGCTCAGGGGCTATATCAAGAAAT	1020
Qy	1021	CAGGAATTCATACGAATTAATTAACGTACAGATATGAGATTTACCTTGAAGAGA	1080
Dp	1021	CAGGAATTCATACGAATTAATTAACGTACAGATATGAGATTTACCTTGAAGAGA	1080
Qy	1081	AGGATTAAGTTGCTAATCCTGATATACAGGACCTAATTTATTCCTCAGGTTAAACA	1140
Dp	1081	AGGATTAAGTTGCTAATCCTGATATACAGGACCTAATTTATTCCTCAGGTTAAACA	1140
Qy	1141	ATGGAAGAACCTATATTTATGATCAAAAAAGAGATCATTAATCTTCGACTGACA	1200
Dp	1141	ATGGAAGAACCTATATTTATGATCAAAAAAGAGATCATTAATCTTCGACTGACA	1200
Qy	1201	TTAACCAAGGGGCGGATGCTTATTTATTTGAGGAGTAAATTTACAGATCTCAAAATTC	1260
Dp	1201	TTAACCAAGGGGCGGATGCTTATTTATTTGAGGAGTAAATTTACAGATCTCAAAATTC	1260
Qy	1261	ACCAAACTTGGCAAGAGCTGGCAATACATGTAAGTAAATAGCAACCGTTACTGGAAAG	1320
Dp	1261	ACCAAACTTGGCAAGAGCTGGCAATACATGTAAGTAAATAGCAACCGTTACTGGAAAG	1320
Qy	1321	TAAATGGCGTGGAACTGATGACCTTCTTAAATTTGTTAAGAGACATTTGACGTTCAAG	1380
Dp	1321	TAAATGGCGTGGAACTGATGACCTTCTTAAATTTGTTAAGAGACATTTGACGTTCAAG	1380
Qy	1381	CGAAAGGCGGAAATTAAGGTTTTCATGAGGTGAGGATGCTAAAGTCAATTTGGACAGC	1440
Dp	1381	CGAAAGGCGGAAATTAAGGTTTTCATGAGGTGAGGATGCTAAAGTCAATTTGGACAGC	1440
Qy	1441	AGGCGACGATCAAGCACAACACCTTTAGTGAATTTGCTTGGTACGGCGAGAG	1500

Db 1441 AGGAGAGATGAGGAGCAAAACAGCTTTAGTAAATGGCTTGGTAAAGGCAAG 1500
Qy 1501 GGAAGTGTCAATTAACATGATTAACATTTGATACGATTAATTTATTTGCTTC 1560
Db 1501 GGAAGTGTCAATTAACATGATTAACATTTGATACGATTAATTTATTTATTTGCTTC 1560
Qy 1561 GTGGTGGTCTTGAATCTTAAACGGGCAATTCATTAACCTTTAAAGTATCAAAATACG 1620
Db 1561 GTGGTGGTCTTGAATCTTAAACGGGCAATTCATTAACCTTTAAAGTATCAAAATACG 1620
Qy 1621 AGGAGGAGGCAATGATTTGGAACCAATTAACCTTAACCTTTAAAGTATCAAAATACG 1680
Db 1621 AGGAGGAGGCAATGATTTGGAACCAATTAACCTTAACCTTTAAAGTATCAAAATACG 1680
Qy 1681 GGAACGAAGATTTGTTACCTTAAGGAAATTAATTAATTAACCTTTAAAGTATCAAAAG 1740
Db 1681 GGAACGAAGATTTGTTACCTTAAGGAAATTAATTAATTAACCTTTAAAGTATCAAAAG 1740
Qy 1741 AAATGGCTTAAACAGGTTGGTGGCAAAACAGATTAATTAACCAATGGGCAATTA 1800
Db 1741 AAATGGCTTAAACAGGTTGGTGGCAAAACAGATTAATTAACCAATGGGCAATTA 1800
Qy 1801 ACCATTTTAAACCAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 ACCATTTTAAACCAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Qy 1861 TAAAGGCGATTTATCCCAACAAAGGTTAACTTTTCAAGGCGTAAAGCAACGCG 1920
Db 1861 TAAAGGCGATTTATCCCAACAAAGGTTAACTTTTCAAGGCGTAAAGCAACGCG 1920
Qy 1921 AGGCTTCAATCAATTTAAATTAACGTTGGTCAAAATGGAAGTATCAACAAAGCGAA 1980
Db 1921 AGGCTTCAATCAATTTAAATTAACGTTGGTCAAAATGGAAGTATCAACAAAGCGAA 1980
Qy 1981 TTGTGTGGATCAAGATTTGATCAACGTTCAATTTAAAGCTGAAATCTTCAATTAAG 2040
Db 1981 TTGTGTGGATCAAGATTTGATCAACGTTCAATTTAAAGCTGAAATCTTCAATTAAG 2040
Qy 2041 GGGAGAGGCGGTTGTTCTCGCAATGTTCTTCAATTTAGGAAATTTGAGCAAGTCA 2100
Db 2041 GGGAGAGGCGGTTGTTCTCGCAATGTTCTTCAATTTAGGAAATTTGAGCAAGTCA 2100
Qy 2101 AATATGCAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 AATATGCAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Qy 2161 CAGATGCAAGATTTAAAGCTTGTCAAAAGTGAATTTAAACCAATGATTTAA 2220
Db 2161 CAGATGCAAGATTTAAAGCTTGTCAAAAGTGAATTTAAACCAATGATTTAA 2220
Qy 2221 ATTCTATACCAAAACCAATCAATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 ATTCTATACCAAAACCAATCAATGATGATGATGATGATGATGATGATGATGATGAT 2280
Qy 2281 ATGTTAAAGGTTTGAAGAACTTAATGCAATGATGATGATGATGATGATGATGATGAT 2340
Db 2281 ATGTTAAAGGTTTGAAGAACTTAATGCAATGATGATGATGATGATGATGATGATGAT 2340
Qy 2341 CATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2400
Db 2341 CATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2400
Qy 2401 CGGTGATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2460
Db 2401 CGGTGATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2460
Qy 2461 TAAAGGCGATTTTGGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2520
Db 2461 TAAAGGCGATTTTGGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2520
Qy 2521 AAAATGAGCTTGAAGATGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580
Db 2521 AAAATGAGCTTGAAGATGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580
Qy 2581 GTACGATCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2640
Db 2581 GTACGATCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2640
Qy 2641 CATTAAGGAGGAAACAGGCAATGCGAATGCGAATGCGAATGCGAATGCGAATGCGAATGCG 2700
Db 2641 CATTAAGGAGGAAACAGGCAATGCGAATGCGAATGCGAATGCGAATGCGAATGCGAATGCG 2700
Qy 2701 ATGTTAAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2760
Db 2701 ATGTTAAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2760
Qy 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
Db 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
Qy 2821 CAGGCAAGAACCCGGAATCCTTGAAGCAATTAATTAATTAATTAATTAATTAATTAATTA 2880
Db 2821 CAGGCAAGAACCCGGAATCCTTGAAGCAATTAATTAATTAATTAATTAATTAATTAATTA 2880
Qy 2881 CGTTATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
Db 2881 CGTTATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
Qy 2941 GTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000
Db 2941 GTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000
Qy 3001 AATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
Db 3001 AATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
Qy 3061 TTGACCGATGCTTAAACCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3120
Db 3061 TTGACCGATGCTTAAACCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3120
Qy 3121 GAGGAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
Db 3121 GAGGAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
Qy 3181 AAGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Db 3181 AAGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Qy 3241 AAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Db 3241 AAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Qy 3301 TTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Db 3301 TTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Qy 3361 AAAAGCAAGCAACCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
Db 3361 AAAAGCAAGCAACCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
Qy 3421 TATCTGCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Db 3421 TATCTGCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Qy 3481 ATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
Db 3481 ATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
Qy 3541 ATGAGGTTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3600
Db 3541 ATGAGGTTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3600
Qy 3601 CTTAGCTTAATGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3660
Db 3601 CTTAGCTTAATGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3660

QY 3661 ATGAACAGTTAAATATCAGCAGATTACATGATGTCGGTTTGGCCATATCAT 3720
 DB 3661 ATGAACAGTTAAATATCAGCAGATTACATGATGTCGGTTTGGCCATATCAT 3720
 QY 3721 GGGGCGATTACATTTGGTGTAAACGTGGAAACGGGAATCAGTCGCAATTAATGCGTG 3780
 DB 3721 GGGGCGATTACATTTGGTGTAAACGTGGAAACGGGAATCAGTCGCAATTAATGCGTG 3780
 QY 3781 AAGAACAAAGCGAAATATCATGCAAAACGATTAATTAATGCGCAATTCGCAATTC 3840
 DB 3781 AAGAACAAAGCGAAATATCATGCAAAACGATTAATTAATGCGCAATTCGCAATTC 3840
 QY 3841 AGTTCCCTTTAGGGCAATTTGGGCAATTCAGCTTAATTTGAGTAATGCTATTTATG 3900
 DB 3841 AGTTCCCTTTAGGGCAATTTGGGCAATTCAGCTTAATTTGAGTAATGCTATTTATG 3900
 QY 3901 AACGTAAAAATATCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3960
 DB 3901 AACGTAAAAATATCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3960
 QY 3961 GGTAAATGCTGAGATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
 DB 3961 GGTAAATGCTGAGATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
 QY 4021 AACCTTAATTTCTGCAATTAATGATGATGATGATGATGATGATGATGATGATGAT 4080
 DB 4021 AACCTTAATTTCTGCAATTAATGATGATGATGATGATGATGATGATGATGATGAT 4080
 QY 4081 ATCTCAGCGTGTGCAACCAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140
 DB 4081 ATCTCAGCGTGTGCAACCAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140
 QY 4141 CAGAAATTTACATTTCCCAATTTCCGCTTTTACGAAATTCAGAGTTCACAGCTG 4200
 DB 4141 CAGAAATTTACATTTCCCAATTTCCGCTTTTACGAAATTCAGAGTTCACAGCTG 4200
 QY 4201 GCMAACAGCAAAATGCGGCGTGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 4260
 DB 4201 GCMAACAGCAAAATGCGGCGTGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 4260
 QY 4261 ATCGTTATTTGATTAACAAAGTGGTCAATTCAGATCCACCTTTTATTCATTAAT 4319
 DB 4261 ATCGTTATTTGATTAACAAAGTGGTCAATTCAGATCCACCTTTTATTCATTAAT 4319

RESULT 3

US-10-680-505-1

Sequence 1, Application US/10080505

Patent No. 667948

GENERAL INFORMATION:

APPLICANT: St. Gene, Joseph W.

TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS

CURRENT APPLICATION NUMBER: US/10/080,505

PRIORITY FILING DATE: 2002-02-22

PRIORITY FILING DATE: 1994-10-25

PRIORITY FILING DATE: 2001-04-20

NUMBER OF SEQ ID NOS: 58

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 4319

TYPE: DNA

ORGANISM: Haemophilus influenzae

FEATURE:

NAME/KEY: CDS

LOCATION: (60)..(4241)

OTHER INFORMATION:

US-10-080-505-1

Query Match 100.0%; Score 4319; DB 4; Length 4319;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAGTGTGTTACTAGATTTTAAATGCAAAATTAATTAATTAATTAATTAATTAATTAAT 60
 DB 1 TCATAGTGTGTTACTAGATTTTAAATGCAAAATTAATTAATTAATTAATTAATTAATTAAT 60
 QY 61 TGAATAAAAGCTGATTTGCTTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
 DB 61 TGAATAAAAGCTGATTTGCTTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
 QY 121 CGAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 121 CGAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 181 CCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
 DB 181 CCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
 QY 241 GCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
 DB 241 GCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
 QY 301 GTAACGGGTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 301 GTAACGGGTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 361 GATATCAATGATGATTTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 DB 361 GATATCAATGATGATTTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 421 ATAAAGTTGTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 DB 421 ATAAAGTTGTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 QY 481 ATAAAGTTGTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
 DB 481 ATAAAGTTGTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
 QY 541 ATAAAGTTGTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 DB 541 ATAAAGTTGTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 QY 601 CTGAGAGGAGATTTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
 DB 601 CTGAGAGGAGATTTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
 QY 661 ATTAATCTGACAGCTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
 DB 661 ATTAATCTGACAGCTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
 QY 721 TGGAGAGGAGATTTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
 DB 721 TGGAGAGGAGATTTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
 QY 781 GGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 781 GGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 841 GGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 841 GGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 901 AATCTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
 DB 901 AATCTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
 QY 961 GTAATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
 DB 961 GTAATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
 QY 1021 CAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080

Db 1021 CAGGATACCATCAGAAATTAATAATAGTAGCAAAATAGATTACCTTTGAAGAGA 1080
 Qy 1081 AGAATTAAGTTCAATATCTAGATATAGAGGCTTAATATTATCTCCAGCTTAAACA 1140
 Db 1081 AGAATTAAGTTCAATATCTAGATATAGAGGCTTAATATTATCTCCAGCTTAAACA 1140
 Qy 1141 ATGGAAGAAACGCTATATTTATGATCAAAAACAGAGATCTATATCTCGATCTGACA 1200
 Db 1141 ATGGAAGAAACGCTATATTTATGATCAAAAACAGAGATCTATATCTCGATCTGACA 1200
 Qy 1201 TTAACCAAGGAGCGAGTGTCTTATTTAGAGGTATTTTACGATCTCCAAATCTA 1260
 Db 1201 TTAACCAAGGAGCGAGTGTCTTATTTAGAGGTATTTTACGATCTCCAAATCTA 1260
 Qy 1261 ACCAACTTGGCAAGGAGTGGCATATAGTAAGAAAATAGACCGTTACTTGGAAAG 1320
 Db 1261 ACCAACTTGGCAAGGAGTGGCATATAGTAAGAAAATAGACCGTTACTTGGAAAG 1320
 Qy 1321 TAAATGGCGTGAACATGATGACCTTCTAAATTTGTAAGAGACATTTGACGTTCAAG 1380
 Db 1321 TAAATGGCGTGAACATGATGACCTTCTAAATTTGTAAGAGACATTTGACGTTCAAG 1380
 Qy 1381 CCAAGAGGAGAAATTAAGGTTGATCGACGGTGAAGGAGATGTAAAGTCTTTGGAGAGC 1440
 Db 1381 CCAAGAGGAGAAATTAAGGTTGATCGACGGTGAAGGAGATGTAAAGTCTTTGGAGAGC 1440
 Qy 1441 AGGAGAGCATCAAGGCAACAAACAAGCCTTTAGTAATTTGCTTGGTTAGCGGACAG 1500
 Db 1441 AGGAGAGCATCAAGGCAACAAACAAGCCTTTAGTAATTTGCTTGGTTAGCGGACAG 1500
 Qy 1501 GGAATCTTCAATTAACGATGATTAACATTTGATACGATTAATTTATTTGGCTTTC 1560
 Db 1501 GGAATCTTCAATTAACGATGATTAACATTTGATACGATTAATTTATTTGGCTTTC 1560
 Qy 1561 GTGGTGGTGGCTTAAATCTTAAAGGAGATTTCACTTAAACCTTTAAACGATTAACG 1620
 Db 1561 GTGGTGGTGGCTTAAATCTTAAAGGAGATTTCACTTAAACCTTTAAACGATTAACG 1620
 Qy 1621 ACGAGGGGCAATGATGTGAAACCATATATCAACTCAAGCGGTAATGTCACTATCTG 1680
 Db 1621 ACGAGGGGCAATGATGTGAAACCATATATCAACTCAAGCGGTAATGTCACTATCTG 1680
 Qy 1681 GAAACGAAAGCATTTGTTCACTATATGAAATATATTTATTAACCTTGAATTAACGAAAG 1740
 Db 1681 GAAACGAAAGCATTTGTTCACTATATGAAATATATTTATTAACCTTGAATTAACGAAAG 1740
 Qy 1741 AAATTTGCTTCAACAGGTTGGTTTGGGAAACAGATTAATAATAACCAATGGGCGATTAA 1800
 Db 1741 AAATTTGCTTCAACAGGTTGGTTTGGGAAACAGATTAATAATAATAACCAATGGGCGATTAA 1800
 Qy 1801 ACCTTATTTAATTAACCAACAGAGAGATGTACTTGTCTTCACTTCAAGTGGTGAATTT 1860
 Db 1801 ACCTTATTTAATTAACCAACAGAGAGATGTACTTGTCTTCACTTCAAGTGGTGAATTT 1860
 Qy 1861 TAAAGGCGATATTACCAACCAAAAGGTAAACTATTTTTCAGCGGTAGACGACACGCG 1920
 Db 1861 TAAAGGCGATATTACCAACCAAAAGGTAAACTATTTTTCAGCGGTAGACGACACGCG 1920
 Qy 1921 ACAGCTTCAATCAATTAATAATAAGTTGGTCAAGAAATGAGAGGTATCAACAGGCGTAA 1980
 Db 1921 ACAGCTTCAATCAATTAATAATAAGTTGGTCAAGAAATGAGAGGTATCAACAGGCGTAA 1980
 Qy 1981 TTGTGTGGATCAAGATGATGATCAACCGTATACCTTTAAAGCTGAAACCTTCCAAATTAAG 2040
 Db 1981 TTGTGTGGATCAAGATGATGATCAACCGTATACCTTTAAAGCTGAAACCTTCCAAATTAAG 2040
 Qy 2041 GGGAGAGTGGGTGGTTTCTGCAATGTCTTCAATTTGAGGAAATTTGACAGTCAACA 2100
 Db 2041 GGGAGAGTGGGTGGTTTCTGCAATGTCTTCAATTTGAGGAAATTTGACAGTCAACA 2100
 Qy 2101 AATAATGAAATGCAATTTGT 2160
 Db 2101 AATAATGCAATGCAATTTGT 2160

Qy 2161 CAGATGAGACGATTAACGACTTGTCAAAAAGTGAATTAACGATACCAAAAGTATTA 2220
 Db 2161 CAGATGAGACGATTAACGACTTGTCAAAAAGTGAATTAACGATACCAAAAGTATTA 2220
 Qy 2221 ATTCTATACCAAAACCAAAATCAATGCTTATTAATTAATTAATTAATTAATTAATTAATTA 2280
 Db 2221 ATTCTATACCAAAACCAAAATCAATGCTTATTAATTAATTAATTAATTAATTAATTAATTA 2280
 Qy 2281 AATGTTAAGTTTGAAGAACTTAATGAGATGTCACTTTAACAATCAACGCAATTTA 2340
 Db 2281 AATGTTAAGTTTGAAGAACTTAATGAGATGTCACTTTAACAATCAACGCAATTTA 2340
 Qy 2341 CATTAGCAACATGACCAACCAAAATATGCAATTTGCACTTTCCGACATTTCACTGCA 2400
 Db 2341 CATTAGCAACATGACCAACCAAAATATGCAATTTGCACTTTCCGACATTTCACTGCA 2400
 Qy 2401 CGGTGATTAAGCAACCTTGAAGGTAATGTGATTAACGATTAACGATTAACGATTAACG 2460
 Db 2401 CGGTGATTAAGCAACCTTGAAGGTAATGTGATTAACGATTAACGATTAACGATTAACG 2460
 Qy 2461 TAAAAACAGCAATTTTTCGACCAAAATTCAGGAGCAAAAGGCAACAGTGAAGTGG 2520
 Db 2461 TAAAAACAGCAATTTTTCGACCAAAATTCAGGAGCAAAAGGCAACAGTGAAGTGG 2520
 Qy 2521 AAAATGCACTTGGACAAATGCTAGAGATTAATGATTAACGATTAACGATTAACGATTAAC 2580
 Db 2521 AAAATGCACTTGGACAAATGCTAGAGATTAATGATTAACGATTAACGATTAACGATTAAC 2580
 Qy 2581 GTACGATCACTTAATTAACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAAC 2640
 Db 2581 GTACGATCACTTAATTAACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAAC 2640
 Qy 2641 CATTAGAGGAGAAACCAAGGCAACATGAGGAGAAAGTGTTCACCACTTGAAGCTTA 2700
 Db 2641 CATTAGAGGAGAAACCAAGGCAACATGAGGAGAAAGTGTTCACCACTTGAAGCTTA 2700
 Qy 2701 ATGCTAAATGAGTGGGCAAGGCAATTCGAAATTTCACTTCACTTATTTTGGCTATTA 2760
 Db 2701 ATGCTAAATGAGTGGGCAAGGCAATTCGAAATTTCACTTCACTTATTTTGGCTATTA 2760
 Qy 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
 Db 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
 Qy 2821 CAGGCAAAAGCCGAAACCTTTGAGCAATTAATTTTGTGTAAGCAAAATTAATTAAC 2880
 Db 2821 CAGGCAAAAGCCGAAACCTTTGAGCAATTAATTTTGTGTAAGCAAAATTAATTAAC 2880
 Qy 2881 CGTTATCAGATTAAGTCAATTTTCACTTGAAGAAATGACAGCTTATGAGTGAAGTGA 2940
 Db 2881 CGTTATCAGATTAAGTCAATTTTCACTTGAAGAAATGACAGCTTATGAGTGAAGTGA 2940
 Qy 2941 GTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000
 Db 2941 GTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000
 Qy 3001 AATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3060
 Db 3001 AATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3060
 Qy 3061 TTGAACGAGTGTGTAACCAAAACAGGTTGAGGCAAAAGTGGGCTGAGAGGAGCGCA 3120
 Db 3061 TTGAACGAGTGTGTAACCAAAACAGGTTGAGGCAAAAGTGGGCTGAGAGGAGCGCA 3120
 Qy 3121 GAGCAGCGTTTCTGATTAACCTTGTGATCAAAAGCTTTTAAAGCATTAATTAATTAATTA 3180
 Db 3121 GAGCAGCGTTTCTGATTAACCTTGTGATCAAAAGCTTTTAAAGCATTAATTAATTAATTA 3180
 Qy 3181 AAGCTGAATGATGCTGAGAAACCAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 3240
 Db 3181 AAGCTGAATGATGCTGAGAAACCAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 3240

QY 361 GATATACAGATGTTGATTTTGTSCAGAGGAAACAACCCGATCAACATGCTTACTT 420
DB 361 GATATACAGATGTTGATTTTGTSCAGAGGAAACAACCCGATCAACATGCTTACTT 420
QY 421 ATTAAGTTGTAAGCAAAATTAACAACAAAAGATATTTCATCTTATGAGACAT 480
DB 421 ATTAAGTTGTAAGCAAAATTAACAACAAAAGATATTTCATCTTATGAGACAT 480
QY 481 ACCATATCCAGATTACATTAATGTTTACAGAAAGGCGCTCCATGATGATGACTGCA 540
DB 481 ACCATATCCAGATTACATTAATGTTTACAGAAAGGCGCTCCATGATGATGACTGCA 540
QY 541 ATATGATGGCAGTACTTATTCAGATTAAGCAAAATATCCAGACGCTGTGCTATCGCT 600
DB 541 ATATGATGGCAGTACTTATTCAGATTAAGCAAAATATCCAGACGCTGTGCTATCGCT 600
QY 601 CTGAGACGCGAGTTTGGCGAAATGATCAAGACAAAGGCGACCAAGTTGCCGCTCATATC 660
DB 601 CTGAGACGCGAGTTTGGCGAAATGATCAAGACAAAGGCGACCAAGTTGCCGCTCATATC 660
QY 661 ATATCTGACAGCTGCGAATTAACAATCAATCAGCGGAGAGGATATGATATTCGAT 720
DB 661 ATATCTGACAGCTGCGAATTAACAATCAATCAGCGGAGAGGATATGATATTCGAT 720
QY 721 TGGAGGCGCATGTGCTTAAGCGGAGAAATATGCTCAATACCGATTCAGGCTCAAGG 780
DB 721 TGGAGGCGCATGTGCTTAAGCGGAGAAATATGCTCAATACCGATTCAGGCTCAAGG 780
QY 781 GGGACAGTGGTCTCCGATGTTTATATGATGCTGAAAAACAAAATGTTATTAATG 840
DB 781 GGGACAGTGGTCTCCGATGTTTATATGATGCTGAAAAACAAAATGTTATTAATG 840
QY 841 GGAATTAAGCGGAAAGGCAACCTTTGAAGGAAAGAAATGCTTGAATGCTGCA 900
DB 841 GGAATTAAGCGGAAAGGCAACCTTTGAAGGAAAGAAATGCTTGAATGCTGCA 900
QY 901 AATCTTATTTGATGAATTTTCGAAAGAGATTTACATACATCACTTACACCCGACTG 960
DB 901 AATCTTATTTGATGAATTTTCGAAAGAGATTTACATACATCACTTACACCCGACTG 960
QY 961 GTAATGAGTGTACCAATTAATGATGATTAATGATGATGATGATGATGATGAT 1020
DB 961 GTAATGAGTGTACCAATTAATGATGATTAATGATGATGATGATGATGATGAT 1020
QY 1021 CAGGAATACCATCAAGAAATTAATTAAGTGAAGAAATAGAGATTAACCTTGAAGGA 1080
DB 1021 CAGGAATACCATCAAGAAATTAATTAAGTGAAGAAATAGAGATTAACCTTGAAGGA 1080
QY 1081 AGGATTAAGTCTATATCTAGATATGACGCACTAATATTTATTCGACGCTTAAACA 1140
DB 1081 AGGATTAAGTCTATATCTAGATATGACGCACTAATATTTATTCGACGCTTAAACA 1140
QY 1141 ATGAGAAACGCTATATTTATGATCAAAAACAGAGATCACTTATCTCGATGCA 1200
DB 1141 ATGAGAAACGCTATATTTATGATCAAAAACAGAGATCACTTATCTCGATGCA 1200
QY 1201 TTAACGAAGGCGGCTGCTTATTTGAGGGAATTTTACAGATCTCCAAATTTCA 1260
DB 1201 TTAACGAAGGCGGCTGCTTATTTGAGGGAATTTTACAGATCTCCAAATTTCA 1260
QY 1261 ACCAAACCTTGGCAAGAGCTGCGATACATGATAGTAAATAGACCGTTACCTGCAAG 1320
DB 1261 ACCAAACCTTGGCAAGAGCTGCGATACATGATAGTAAATAGACCGTTACCTGCAAG 1320
QY 1321 TAAATGCGGTGGAACATGATGATCTTCTTAAATTTGTAAGAAACATTCGACGCT 1380
DB 1321 TAAATGCGGTGGAACATGATGATCTTCTTAAATTTGTAAGAAACATTCGACGCT 1380
QY 1381 CCAAGGGAAGAAATTAAGGTTGATCAGGCTAGGAGGATGTAAGTCACTTTGAGGACG 1440
DB 1381 CCAAGGGAAGAAATTAAGGTTGATCAGGCTAGGAGGATGTAAGTCACTTTGAGGACG 1440

QY 1441 AGGACAGATCAAGGCAACAAACAGCCTTTAGTAAATTTGGCTGTGTAAGCGGACAG 1500
DB 1441 AGGACAGATCAAGGCAACAAACAGCCTTTAGTAAATTTGGCTGTGTAAGCGGACAG 1500
QY 1501 GGAATGCTCAATTAACGATGATTAACAAATTTGATACCGATTAATTTATTTGGCTTTC 1560
DB 1501 GGAATGCTCAATTAACGATGATTAACAAATTTGATACCGATTAATTTATTTGGCTTTC 1560
QY 1561 GTGGGTGTGCTTATGATCTTAAACGCAATCATTAACCTTTAAACGATCAAAATGCG 1620
DB 1561 GTGGGTGTGCTTATGATCTTAAACGCAATCATTAACCTTTAAACGATCAAAATGCG 1620
QY 1621 AGGAGGCGCATATGATGTAACATTAATCACTCAACGCGCTAATGTCACTATTAAG 1680
DB 1621 AGGAGGCGCATATGATGTAACATTAATCACTCAACGCGCTAATGTCACTATTAAG 1680
QY 1681 GGAAGGAAGCATTTGCTACCTAATGGAATTAATTAATTAACCTTATTAAGGAAAG 1740
DB 1681 GGAAGGAAGCATTTGCTACCTAATGGAATTAATTAATTAACCTTATTAAGGAAAG 1740
QY 1741 AAATGCTTACAGCTGTGTTGGCGAAACAGTAAATTAACACATGAGGCACTTA 1800
DB 1741 AAATGCTTACAGCTGTGTTGGCGAAACAGTAAATTAACACATGAGGCACTTA 1800
QY 1801 ACCTTATTTATTAACCAACCAAGATGCTACTTGTCTACTTCAAGTGTGTAAT 1860
DB 1801 ACCTTATTTATTAACCAACCAAGATGCTACTTGTCTACTTCAAGTGTGTAAT 1860
QY 1861 TAAAGGCGATTTTACCGAAACAAAGGTAATTTTTCAGCGGTGACACGCGC 1920
DB 1861 TAAAGGCGATTTTACCGAAACAAAGGTAATTTTTCAGCGGTGACACGCGC 1920
QY 1921 AAGCCTTACATCATTTAAATTAACGTTGTCAGAAATGGAAGGTTACACAGGCGAA 1980
DB 1921 AAGCCTTACATCATTTAAATTAACGTTGTCAGAAATGGAAGGTTACACAGGCGAA 1980
QY 1981 TTGTGTGGAATCAGATTTGATCAACGCTACATTTAAAGCTGAAATCTTCAATTAAG 2040
DB 1981 TTGTGTGGAATCAGATTTGATCAACGCTACATTTAAAGCTGAAATCTTCAATTAAG 2040
QY 2041 GGGAGAGTGGGTGCTTCTGCAATGTTTCTCAATGAGGAAATGGAAGCTGAGCA 2100
DB 2041 GGGAGAGTGGGTGCTTCTGCAATGTTTCTCAATGAGGAAATGGAAGCTGAGCA 2100
QY 2101 ATATGCAAAACGCACTTTGCTGTGCGCAATCAACAAATCACTTGCACGCGTT 2160
DB 2101 ATATGCAAAACGCACTTTGCTGTGCGCAATCAACAAATCACTTGCACGCGTT 2160
QY 2161 CAGATGGAAGATTAAGCACTTGTCAAAAAGTGAATTAACCGATTCAAAGTTATTA 2220
DB 2161 CAGATGGAAGATTAAGCACTTGTGTCAAAAAGTGAATTAACCGATTCAAAGTTATTA 2220
QY 2221 ATTCTTACCAAAACCAATCAATGAGCTGCTATTAATTAATGATTAAGCAAGGCGCA 2280
DB 2221 ATTCTTACCAAAACCAATCAATGAGCTGCTATTAATTAATGATTAAGCAAGGCGCA 2280
QY 2281 ATGTTAAGGTTTACCAAACTTAATGCAATGTCATTTAAACAAATCAAGCCAAATTA 2340
DB 2281 ATGTTAAGGTTTACCAAACTTAATGCAATGTCATTTAAACAAATCAAGCCAAATTA 2340
QY 2341 CATTAAGCAACATGCAACCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2400
DB 2341 CATTAAGCAACATGCAACCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2400
QY 2401 CGGTGATTAAGCAACTGTAAGGTAAGGCAATTAAGGATCACTTAAATTTCTT 2460
DB 2401 CGGTGATTAAGCAACTGTAAGGTAAGGCAATTAAGGATCACTTAAATTTCTT 2460
QY 2461 TAAAGGAGCATTTTTCGACCAAAATTCAGGAGCAAGGCAACAGTACGTTG 2520
DB 2461 TAAAGGAGCATTTTTCGACCAAAATTCAGGAGCAAGGCAACAGTACGTTG 2520
QY 2521 AAAATGCACTTGAACATGCTTACGATTAATGCAAGATTTAAAGCTTAATTAACA 2580

```

Db 2521 AAAATGCACTGGAATAGCTACGACTACATGCAAACTTAAACCTTAATACCA 2590
Qy 2581 GTAGATCAAGTTAAATCACTATTAGCTAGCTCAACCAATACGCAAGTCCGCTT 2640
Db 2581 GTAGATCAAGTTAAATCACTATTAGCTAGCTCAACCAATACGCAAGTCCGCTT 2640
Qy 2641 CATTAAGACGGAACCAACGCAATCGGCAAGCATGCTTCAACATTAAGACATTA 2700
Db 2641 CATTAAGACGGAACCAACGCAATCGGCAAGCATGCTTCAACATTAAGACATTA 2700
Qy 2701 ATGCTAATTAAGTGGGCAAGGCAATTCATTTCTTATTTTGGCTATTA 2760
Db 2701 ATGCTAATTAAGTGGGCAAGGCAATTCATTTCTTATTTTGGCTATTA 2760
Qy 2761 GCGATTAATTAATTAATCAATGACGCTGAGGCGATTAATTAATTAATTAAT 2820
Db 2761 GCGATTAATTAATTAATCAATGACGCTGAGGCGATTAATTAATTAATTAAT 2820
Qy 2821 CAGGCAAGAAACCGAAACCTTGAGCAATTAATTTGGTGAAGCAAGATTAATCA 2880
Db 2821 CAGGCAAGAAACCGAAACCTTGAGCAATTAATTTGGTGAAGCAAGATTAATCA 2880
Qy 2881 CGTTATCAAGTAAGCTCAATTTACTTAAGAAATGACAGCTTGAAGCAATTA 2940
Db 2881 CGTTATCAAGTAAGCTCAATTTACTTAAGAAATGACAGCTTGAAGCAATTA 2940
Qy 2941 GTTATTAATTAAGTGAAGTATGAGGCAATTCGCTGATTAATTAATTAATTA 3000
Db 2941 GTTATTAATTAAGTGAAGTATGAGGCAATTCGCTGATTAATTAATTAATTA 3000
Qy 3001 AATTGCAATTAATTAATTAAGTGAAGTATGAGGCAATTCGCTGATTAATTA 3060
Db 3001 AATTGCAATTAATTAATTAAGTGAAGTATGAGGCAATTCGCTGATTAATTA 3060
Qy 3061 TTGAACGCACTGCTAAACCAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3120
Db 3061 TTGAACGCACTGCTAAACCAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3120
Qy 3121 GAGCAGCTTTCCTGATACCTGCTGATTAATTAATTAATTAATTAATTAATTA 3180
Db 3121 GAGCAGCTTTCCTGATACCTGCTGATTAATTAATTAATTAATTAATTAATTA 3180
Qy 3181 AAGCTGAATGATGCTGTAACCAACAAAGTGAAGTGAAGTGAAGTGAAGTGA 3240
Db 3181 AAGCTGAATGATGCTGTAACCAACAAAGTGAAGTGAAGTGAAGTGAAGTGA 3240
Qy 3241 AAAAGAGAGTGTTCCTGATACCTGCTGATTAATTAATTAATTAATTAATTA 3300
Db 3241 AAAAGAGAGTGTTCCTGATACCTGCTGATTAATTAATTAATTAATTAATTA 3300
Qy 3301 TTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3360
Db 3301 TTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3360
Qy 3361 AAAAACAACCAACCAACCAACCTTGAATGACGCTTGAATGACGCTTGAAT 3420
Db 3361 AAAAACAACCAACCAACCAACCTTGAATGACGCTTGAATGACGCTTGAAT 3420
Qy 3421 TATTCGACAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Db 3421 TATTCGACAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Qy 3481 ATCAAGACAAATTCGCGTGTGAACAAATTAATTAATTAATTAATTAATTA 3540
Db 3481 ATCAAGACAAATTCGCGTGTGAACAAATTAATTAATTAATTAATTAATTA 3540
Qy 3541 ATGCGTTTCGCTTATTAAGCAAGAAACCACTTACGCAATTTGGGTCGAAAG 3600
Db 3541 ATGCGTTTCGCTTATTAAGCAAGAAACCACTTACGCAATTTGGGTCGAAAG 3600
Qy 3601 CTTAGCAATTAAGCAATTTGGGTCGCTTTCGATTAAGCTTCAATTAATTA 3660

```

```

Db 3601 CTTAGCAATTAAGCAATTTGGGTCGCTTTCGATTAAGCTTCAATTAATTA 3660
Qy 3661 ATGAAACAGGTTAAATCAAGCAATTAATTAATTAATTAATTAATTAATTA 3720
Db 3661 ATGAAACAGGTTAAATCAAGCAATTAATTAATTAATTAATTAATTAATTA 3720
Qy 3721 GGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3780
Db 3721 GGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3780
Qy 3781 AAGAAACAAACCAACCAACCACTTGAATGACGCTTGAATGACGCTTGAAT 3840
Db 3781 AAGAAACAAACCAACCAACCACTTGAATGACGCTTGAATGACGCTTGAAT 3840
Qy 3841 AGTTCCGTTTGAAGCAATTTGGGTCGCTTTCGATTAAGCTTCAATTAATTA 3900
Db 3841 AGTTCCGTTTGAAGCAATTTGGGTCGCTTTCGATTAAGCTTCAATTAATTA 3900
Qy 3901 AAGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3960
Db 3901 AAGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3960
Qy 3961 GCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4020
Db 3961 GCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4020
Qy 4021 AGCTTATTTCTTCGCTCAATTAATTAATTAATTAATTAATTAATTAATTA 4080
Db 4021 AGCTTATTTCTTCGCTCAATTAATTAATTAATTAATTAATTAATTAATTA 4080
Qy 4081 ATCTCAGGTTGCAACCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4140
Db 4081 ATCTCAGGTTGCAACCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4140
Qy 4141 CAGAAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4200
Db 4141 CAGAAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4200
Qy 4201 GCAACAGCAAAATGAGGCGTGAATTAATTAATTAATTAATTAATTAATTA 4260
Db 4201 GCAACAGCAAAATGAGGCGTGAATTAATTAATTAATTAATTAATTAATTA 4260
Qy 4261 ATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4320
Db 4261 ATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4320

```

RESULT 5
 US-10-080-505-8
 : Sequence 8, Application US/10080505
 : Patent No. 6676948
 : GENERAL INFORMATION:
 : APPLICANT: St. Gene, Joseph W.
 : TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 : FILE REFERENCE: A-59941-1/REF/DCF/DEK
 : CURRENT APPLICATION NUMBER: US/10/080,505
 : PRIOR APPLICATION NUMBER: US 08/296,791
 : PRIOR FILING DATE: 1994-10-25
 : PRIOR APPLICATION NUMBER: US 09/839,996
 : PRIOR FILING DATE: 2001-04-20
 : NUMBER OF SEQ ID NOS: 58
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 8
 : LENGTH: 4305
 : TYPE: DNA
 : ORGANISM: Haemophilus influenzae
 : FEATURE:
 : NAME/KEY: misc feature
 : LOCATION: (1702)..(1702)
 : OTHER INFORMATION: "r" at position 1702 can be any base.
 : NAME/KEY: CDS
 : LOCATION: (1)..(4305)

OTHER INFORMATION:

US-10-080-505-8

Query Match 73.6%; Score 3180.8; DB 4; Length 4305;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 3704; Conservative 0; Mismatches 463; Indels 156; Gaps 9;

QY 60 ATGAAAAAACTGATTTCTGCTTAATTTTAAACCGCTGATTCATTAGAGATGA 119
 Db 1 ATGAAAAAACTGATTTCTGCTTAATTTTAAACCGCTGATTCATTAGAGATGA 60
 QY 120 TCGAACGGTGGCTGGTCACTTAATTTGGATTGATTCGAAATTAATGATTTT 179
 Db 61 TCGAACGGTGGCTGGTCACTTAATTTGGATTGATTCGAAATTAATGATTTT 120
 QY 180 GCCGAAATTAAGGGAGGTTCAGAGTGGGGCTCAAAATTAATTAAGGTTATACAAACAA 239
 Db 121 GCCGAAATTAAGGGAGGTTCAGAGTGGGGCTCAAAATTAATTAAGGTTATACAAACAA 180
 QY 240 GGGCAATTAGTTGGCAATCATGACAAAAGCCCGATGATTAATTTCTGATAGTCA 299
 Db 181 GGGCAATTAGTTGGCAATCATGACAAAAGCCCGATGATTAATTTCTGATAGTTC 240
 QY 300 CTTAAGCGCTGGCAGCCTTGGTGAATCATATATTGGAGCTGGCAGATACGTA 359
 Db 241 AGAATGAGATTGCTGCTTAGTAGAGGATCAGTATATTGGATGGCAGATACGTA 300
 QY 360 GGATATACAGATGTTGATTTGGTGCAGAGGGAAACACCCCGATGACATGGTTTACT 419
 Db 301 GGCTATACCAATGATGATTTGGTGCAGAGGAAACAAATTCCTGATACATGGTTTACT 360
 QY 420 TATAAGTTGTAAACGAATTAATCAAAAAGATTAATTAATCTTATAGAGACAT 479
 Db 361 TATAAGTTGTAAACGAATTAATTAATCAAGATGAGAGACACCGCATATAGAGAC 420
 QY 480 TACCATATCCAGATTAATCAATTAATGTTACAGAGGCGCTCAATTGATTAATCTTG 539
 Db 421 TACCATATCCAGATTAATCAATTAATGTTACAGAGGCGCTCAATTGATTAATCTTG 480
 QY 540 AATATGATGGCAGATTAATCAATTAATGATTAATCAATTAATGATTAATGATTAAT 599
 Db 481 CATATGATGGCAGATTAATCAATTAATGATTAATCAATTAATGATTAATGATTAAT 540
 QY 600 TCTGACGGCAGATTTTGGCGAATGATCAAGAACAAAGCGACAGATTCGGTGCATAT 659
 Db 541 TCTGACGGCAGATTTTGGCGAATGATCAAGAACAAAGCGATTCATTAATGAGATAT 600
 QY 660 CATATCTGACGGCTGGCAATTAATCAATTAATGATTAATGATTAATGATTAATGAT 719
 Db 601 AATATCTGACGGCTGGCAATTAATCAATTAATGATTAATGATTAATGATTAATGAT 660
 QY 720 TTGGAGGGGAGATTTGTTAAAGGGAGAAATGGTTCATTAATGATTAATGATTAATGAT 779
 Db 661 GTAGAGGGGATTAATTTGCGAAGTTGGCGATTAATGATTAATGATTAATGATTAATGAT 720
 QY 780 GGGGACAGTGGTCTCCAGATTAATTAATGATTAATGATTAATGATTAATGATTAATGAT 839
 Db 721 GGGGACAGTGGTCTCCAGATTAATTAATGATTAATGATTAATGATTAATGATTAATGAT 780
 QY 840 GGGATTTACGGGAGGGAGAACCTTTTGAAGGCAAGAAATGGGTTCAATGGTTGCG 899
 Db 781 GGAATTTACGGGAGGGAGAACCTTTTGAAGGCAAGAAATGGGTTCAATGGTTGCG 840
 QY 900 AATCTTATTTGATGAATTTTCGAAAGATTAATTAATGATTAATGATTAATGATTAATGAT 959
 Db 841 AATCTTATTTGATGAATTTTCGAAAGATTAATTAATGATTAATGATTAATGATTAATGAT 900
 QY 960 GGTAAATGAGTGTACAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1019
 Db 901 GGTAAATGAGTGTATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 960
 QY 1226 TCGGAATACCAATGAATTAATTAATGATTAATGATTAATGATTAATGATTAATGAT 1079

Db 961 ATTGAAAAAC-----TCGTGAAATTAACCTTTAGCGAACAAATTAATTAATGAAGAT 1014
 QY 1080 AAGATTAAGTTCATTAATCTTAATTAATGAGAGCACTTAATTTATTCACAGTTTAAAC 1139
 Db 1015 AAGATTAAGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1074
 QY 1140 AATGAGAAAGGCTATATTTATTAATGATCAAAACAAAGATTAATTAATTAATTAATTAAT 1199
 Db 1075 AATGAGAAAGGCTATATTTATTAATGATCAAAACAAAGATTAATTAATTAATTAATTAAT 1134
 QY 1200 AATTAACAAAGGCGGGTGTCTTAATTTGAGGATTAATTTAGCTATCCAAATTC 1259
 Db 1135 ATGATCAAGGTGAGCGCGTGTCTTAATTTGAGGATTAATTTAGCTATCCAAATTC 1194
 QY 1260 AACCAACTTGGCAAGAGCTGGCAATTAATTAATGAGAAATTAATTAATTAATTAATTAAT 1319
 Db 1195 AATGAAAGGTGAAAGGTGCGGCAATTAATTAATGAGAAATTAATTAATTAATTAATTAAT 1254
 QY 1320 GTAAATGGGTGAAACATGATGACCTTTCTAAATTTGGTAAAGAACTTGACAGTTCAA 1379
 Db 1255 GTAAATGGGTGAAACATGATGACCTTTCTAAATTTGGTAAAGAACTTGACAGTTCAA 1314
 QY 1380 GCCAAAGGGGAAATTAATGATTAATTAATGAGGATTAATTAATTAATTAATTAATTAAT 1439
 Db 1315 GCCAAAGGGGAAATTAATGATTAATTAATGAGGATTAATTAATTAATTAATTAATTAAT 1374
 QY 1440 CAGGACAGCATCAAGGCAACAAACAGCTTTAGTGAATTTGGTGGTAGCGGAG 1499
 Db 1375 CAGGACAGCATCAAGGCAACAAACAGCTTTAGTGAATTTGGTGGTAGCGGAG 1434
 QY 1500 GGGACGTGTTCAATTAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1559
 Db 1435 GGGACGTGTTCAATTAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1494
 QY 1560 CGTGGTGGTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1619
 Db 1495 CGTGGTGGTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1554
 QY 1620 GACGAGGGGGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1679
 Db 1555 GACGAGGGGGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1614
 QY 1680 GGGAGCAAGAACAT--TGTTCTACTTAATGAAATTAATTAATTAATTAATTAATTAATTAAT 1736
 Db 1615 GGGAGCAAGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1674
 QY 1737 AAGAAATTTGCTTAACAGCTTTGGTGGAAACGATTAATTAATTAATTAATTAATTAATTAAT 1796
 Db 1675 AAGAAATTTGCTTAACAGCTTTGGTGGAAACGATTAATTAATTAATTAATTAATTAATTAAT 1734
 QY 1797 TTAACCTTATTAATTAACCAACAGAGATGTAATTAATTAATTAATTAATTAATTAATTAAT 1856
 Db 1735 TTAACCTTATTAATTAACCAACAGAGATGTAATTAATTAATTAATTAATTAATTAATTAAT 1794
 QY 1857 AATTTAAAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1916
 Db 1795 AATTTAAAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1854
 QY 1917 CCGCACGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1976
 Db 1855 CCGCACGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1914
 QY 1977 GAAATTTGTTGGATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2036
 Db 1915 GAAATTTGTTGGATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1974
 QY 2037 AAGGCGGAAGTGGGTGTTTCTGCAATGTTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 2096
 Db 1975 AAGGCGGAAGTGGGTGTTTCTGCAATGTTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 2034
 QY 2097 AGCAATTAATGAAATGCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2156
 Db 2035 AGCAATTAATGAAATGCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2094

QY 2157 CGTTCAGATTGACAGAGATTACAGACTTGTCAAAAAGTGATCTTAACCATACAAAGTT 2216
DB 2095 CGTTCAGATTGACAGAGATTACAGACTTGTCAAAAAGTGATCTTAACCATACAAAGTT 2154
QY 2217 ATTATTTCTTACCAAAAACAAATCATATGCTCTATTATTAATTAATGATGCAACG 2276
DB 2155 ATTATTTCTTACCAAAAACAAATCATATGCTCTATTATTAATTAATGATGCAACG 2214
QY 2277 GCGAATGTAAAGCTTTGCAAAAATTAAATGCAATGCTCTTAACCATACAGCCAA 2336
DB 2215 GTGAATTTATATGTTTACCAAACTTAATGATGCTCTTAATTAATTAATGATGCAACG 2274
QY 2337 TTTCATTAAAGCAACATGCAACCAATTAATGCAATTAATGCAATTAATGCAATTAATGCA 2396
DB 2275 TTTCATTAAAGCAACATGCAACCAATTAATGCAATTAATGCAATTAATGCAATTAATGCA 2334
QY 2397 ----- 2396
DB 2335 GCAAGGTAAATTAATGCACTTTAATGCGGCAATGCAATTAATGCAATTAATGCAATTAATGCA 2394
QY 2397 -----GCA 2399
DB 2395 ACATTAAAGCAATTAATGCACTTTAATGCGGCAATGCAATTAATGCAATTAATGCAATTAATGCA 2354
QY 2400 ACGGTGATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2459
DB 2455 ACGGTGATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2514
QY 2465 TTTAAAAACAGCAATTTTTCGCAACCAATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2519
DB 2515 TTTAAAAACAGCAATTTTTCGCAACCAATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2514
QY 2520 GAAATTCGACTTGAACCAATGCTTAAGCAATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2579
DB 2575 GAAATTCGACTTGAACCAATGCTTAAGCAATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2534
QY 2580 AGTACATCACTTGAACCAATGCTTAAGCAATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2636
DB 2635 AGTACATCACTTGAACCAATGCTTAAGCAATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2634
QY 2637 CGTTCATTAAAGCAACCAATGCTTAAGCAATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2696
DB 2695 CGTTCATTAAAGCAACCAATGCTTAAGCAATTAATGCACTTTAATGCACTTTAATGCACTTTAATGCA 2754
QY 2697 GTAATGTAAATTAATGCTTAAGCAACCAATGCTTAAGCAATTAATGCACTTTAATGCACTTTAATGCA 2756
DB 2755 GTAATGTAAATTAATGCTTAAGCAACCAATGCTTAAGCAATTAATGCACTTTAATGCACTTTAATGCA 2814
QY 2757 AAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2816
DB 2815 AAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2814
QY 2817 AAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2876
DB 2875 AAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2834
QY 2877 CAACGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2936
DB 2935 CAACGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2934
QY 2937 TTAAGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2996
DB 2995 TTAAGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3054
QY 2997 CAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3056
DB 3055 CAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3114
QY 3057 CAAGTTAAACCACTGCTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3116
DB 3115 CAAGTTAAACCACTGCTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3168

QY 3117 GCGAGAGAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3176
DB 3165 ---AGAGAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3225
QY 3177 AAACAGCTGAATGCTGCTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3236
DB 3226 AAACAGCTGAATGCTGCTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3285
QY 3237 TCAAAAAGAGC-----AGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3284
DB 3286 TCAAAAAGAGCATTGAGAGAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3345
QY 3285 GCAATTAAGAGAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3344
DB 3346 GCAATTAAGAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3405
QY 3345 GCTCAAA---GAGAGAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3401
DB 3406 GCTCAAAAGAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3465
QY 3402 AATAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3461
DB 3466 AATAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3525
QY 3462 TTAAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3521
DB 3526 TTAAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3585
QY 3522 AAAAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3581
DB 3586 AAAAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3642
QY 3582 CAATTTGGGTTGCAAAAGCTTTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3641
DB 3643 CAATTTGGGTTGCAAAAGCTTTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 3702
QY 3642 CGTTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3701
DB 3703 CGTTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3762
QY 3702 GGTTCGCTTGAACCAATGCTTAAGCAACCAATGCTTAAGCAACCAATGCTTAAGCAACCAATGCTTAAG 3761
DB 3763 GGTTCGCTTGAACCAATGCTTAAGCAACCAATGCTTAAGCAACCAATGCTTAAGCAACCAATGCTTAAG 3822
QY 3762 AGTACAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3821
DB 3823 AGTACAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3882
QY 3822 GGTTCGCTTGAACCAATGCTTAAGCAACCAATGCTTAAGCAACCAATGCTTAAGCAACCAATGCTTAAG 3881
DB 3883 GGTTCGCTTGAACCAATGCTTAAGCAACCAATGCTTAAGCAACCAATGCTTAAGCAACCAATGCTTAAG 3942
QY 3882 GTTAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3941
DB 3943 GTTAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4002
QY 3942 CTTAGGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4001
DB 4003 CTTAGGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4062
QY 4002 ACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4061
DB 4063 ACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4122
QY 4062 ACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4121
DB 4123 ACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4182
QY 4122 AAAAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 4181
DB 4183 AAAAGAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 4242
QY 4182 TCTCAAGGTTTCTGATACCTGCTGATACCAAGCTTTAAAGCATTAGAACGC 4241

Db 4243 TCCTCAAGCTTCACTGCGTAAACGAAATGTGGCGTGAATTTGGGCTATGCTGG 4302
 QY 4242 TAA 4244
 Db 4303 TAA 4305

RESULT 6

US-09-557-884-1/C
 : Sequence 1, Application US/09557884
 : Patent No. 6506581

GENERAL INFORMATION:

APPLICANT: Pleischmann et al.
 TITLE OF INVENTION: The Nucleotide sequence of
 the Haemophilus influenzae Rd Genome, Fragments
 thereof, and uses thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville

STATE: MD
 COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:
 NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PE186P3

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 69.9%; Score 3019.8; DB 4; Length 1830121;
 Best Local Similarity 83.0%; Pred. No. 0;

Matches 3641; Conservative 2; Mismatches 650; Indels 93; Gaps 14;

QY 1 TCATATGCTGTTAATA-GTATTTTAAACGAAATTAATCTTAATTAATAAAGCTT 59
 Db 278587 TCATATGCTGTTAATAACGATTTTAAACGAAATTAATCTTAATTAATAAAGCTT 278528
 QY 60 ATGAAAAAACTGATTTGCTTAATTTTAAACGCTGTCATTTCAATTAAGGATAGTA 119
 Db 278527 ATGAAAAAACTGATTTGCTTAATTTTAAACGCTGTCATTTCAATTAAGGATAGTA 278468
 QY 120 TCGAAGCGTGGGCTGGTCACTTAATTTGGGATTGATTCATCAATTAATGCGATTTT 179
 Db 278467 TCGAAGCGTGGGCTGGTCACTTAATTTGGGATTGATTCATCAATTAATGCGATTTT 278408
 QY 180 GCCGAGAAATTAAGGAGGATTCACAGTGGGGCTCAAAATATTAAGGTTATTAAGAAACA 239
 Db 278407 GCCGAGAAATTAAGGAGGATTCACAGTGGGGCTCAAAATATTAAGGTTATTAAGAAACA 278348

QY 240 GGGCAATTAAGTGGCAATCAATGACAAAAGCCCGATGATTTTCTGTAGTGTCA 299
 Db 278347 GGGCAATTAAGTGGCAATCAATGACAAAAGCCCGATGATTTTCTGTAGTGTGTG 278288
 QY 300 CGTACGCGGTGGGAGGCTTTGGTTGAAAATCAATATATTTGGAGGTGGCAATTAAGTA 359
 Db 278287 CGTACGCGGTGGGAGGCTTTGGTTGAAAATCAATATATTTGGAGGTGGCAATTAAGTA 278228
 QY 360 GGAATATCAATGTTGATTTTGGTGCAGAGGAAAACCCCGATCAATCGTTTACT 419
 Db 278227 GGAATATCAATGTTGATTTTGGTGCAGAGGAAAACCCCGATCAATCGTTTACT 278168
 QY 420 TATTAATGTTAAACCAATTAATCTACAA---AAAGTATTTTATCTCTTAAGAGGAC 476
 Db 278167 TATTAATGTTAAACCAATTAATCTACAA---AAAGTATTTTATCTCTTAAGAGGAC 278108
 QY 477 GATTACATATTCACAGATTAACCAATTTGTTACAGAACGGCTCAATTGATATGACT 536
 Db 278107 GATTACATATTCACAGATTAACCAATTTGTTACAGAACGGCTCAATTGATATGACT 278048
 QY 537 TCGAATATGATGCGACTACTTATTCAGATGACAAAATATTCAGAAAGCTGTGTATC 596
 Db 278047 TCGAATATGATGCGACTACTTATTCAGATGACAAAATATTCAGAAAGCTGTGTATC 277988
 QY 597 GGCTGTGACGCGAGCTTTTGCGAAATATATCAAGACAAAGCG- 639
 Db 277987 GGCTGTGACGCGAGCTTTTGCGAAATATATCAAGACAAAGCGATATCAATGCTCA 277928
 QY 640 -ACCAAGTTCGCGGTGCATATCTTATCTGACAGCTGGCAATACACAAATCAAGCTGCA 698
 Db 277927 TATTAATGCTGAGGTGCATATCTTATCTGACAGAGTATTCACAAATCAAGCTGCA 277868
 QY 699 GCGAGTAATGATATTTGTTTGGAGGCGATTTGTTAAACGGGAGAAATATGTGCA 758
 Db 277867 GCGAGTAATGATATTTGTTTGGAGGCGATTTGTTAAACGGGAGAAATATGTGCA 277808
 QY 759 TTACCATATGACAGGCTCAAAAGGGGAGAGGTGCTCCAGTGTATTTATTAATGATGCTGA 818
 Db 277807 TTACCATATGACAGGCTCAAAAGGGGAGAGGTGCTCCAGTGTATTTATTAATGATGCTGA 277748
 QY 819 AAACAAAATGTTAATTAATTAATGGAATTTACGGAGGCAACCTTTTGAAGGCAAGAA 878
 Db 277747 AAACAAAATGTTAATTAATTAATGGAATTTACGGAGGCAACCTTTTGAAGGAGGT 277688
 QY 879 AATGGTTTCATTTGTTGCGAAATCTTATTTGTA-----GAATTTTC 923
 Db 277687 AATGGTTTCATTTGTTGCGAAATCTTATTTGTA-----GAATTTTC 277628
 QY 924 GAAGAGATTTACATCAATCACTTAACCCGAGCTGTATGAGGTGTACCAATTAAGT 983
 Db 277627 CTAAGTGTTCACAGGCTATATTCCTCCCATTAATGACATTTATCTTTGTTCAAAAT 277568
 QY 984 GGAATGATATATGTCAGGGGCTATTAACAGAAATACAGAAATCAATCAAGAAATTAAT 1043
 Db 277567 AATGATGATATATGTCAGGAATTAATTAACCTTAAGTAAAGATGAGGCTCAAGCAAAA 277508
 QY 1044 ATTAAGTATGAATATGAGTTTACT-----TTGAAGAGAGAGATTAAGTTCAAT 1094
 Db 277507 TCGAAGTATGAGTATGAGTTTATTTATCAATGTTAAATCAACCGCTTAAGAAACAT 277448
 QY 1095 AATCTATGATATGACGAGCTATATTTATTTTCCAGTTTAAACATGAGAAAGGCTTA 1154
 Db 277447 GTTAAGGACAGCAGAGGCTATATTTATTTTCAACAGAGATGGAATATGAAAAATATAT 277388
 QY 1155 TATTTATGATCAAAAACAGGATCATTAATCTTGCAATCGAATTAACCAAGGGGCG 1214
 Db 277387 TACCTTGGGACCAAGAGAAAGGAACTTTAACATCAAAAATTAATTAACAGGTCGT 277328
 QY 1215 GGGGTCCTTTATTTTGAAGGTTATTTTACATCTGCA-----ATTCACCAAACTTGG 1271
 Db 277327 GGGGTCCTTTATTTTGAAGGTTATTTTGTGTTAAAGGCAAGCAAAAATTAATTAACCTGG 277268
 QY 1272 CAGGAGCTGCGATATATGTAAGTAAATAGCACCGTTACTTGGAAAGTAAATGCGGTC 1331

Db 277267 CAGAGTGGGCGTATCTATGAGCAAGATGCAAGCTGATGAGTGAAGGTTCCAAATCT 277208
Qy 1332 GAACATGATGCACTTCTTAAATTTGGTAAAGAAACATGCACTGCAAGCCAAAGGCA 1331
Db 277207 GAAATGATGCTTATCTTAAATTTGGTAAAGAAACATGCACTGCAAGCCAAAGGCA 277148
Qy 1392 AATTAAGTTGATGATGAGCGATGAGTAAAGTCAATTTGGAGCAGCAGCAGAT 1451
Db 277147 AATTAAGTTGATGAGCGATGAGTAAAGTCAATTTGGAGCAGCAGCAGAT 277088
Qy 1452 CAGGCAACAAACAGCCTTTAGTGAATTTGCTGCTTACGCGCAGAGGACTGTCA 1511
Db 277087 GCGGGTCAAAAACAGCTTTCAAAAGATGGCATGTAAAGCGGTCCAGCAACATGCA 277028
Qy 1512 TTAAGCATGATAAACATTTGATACCGATTAATTTATTTGCGCTTCTGCTGCTGCG 1511
Db 277027 TTAAGCATGATAAACATTTGATACCGATTAATTTATTTGCGCTTCTGCTGCTGCG 276968
Qy 1572 TTGATCTTAAACGCGCATTCATTAACCTTTAAACGATCCAAATACGACGAGGCGCA 1611
Db 276967 TTGATCTTAAACGCGCATTCATTAACCTTTAAACGATCCAAATACGACGAGGCGCG 276928
Qy 1632 ATGATGTGAAACATTAATCACTCAACGCGCTAATGCTACTATTAATGAGAACAAAGC 1631
Db 276907 ATGATGTGAAACATTAATCACTCAACGCGCTAATGCTACTATTAATGAGAACAAAGT 276848
Qy 1692 AT---TGTTACCTTATGAAATATATTAATTAATTAATTAATTAATTAATTAATTA 1748
Db 276847 ATTTACTGCTCATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 276788
Qy 1749 TACAACGCTTGGTTGGCGAAACAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1838
Db 276787 TACAACGCTTGGTTGGCGAAACAGATTAATTAATTAATTAATTAATTAATTAATTAAT 276728
Qy 1809 TTTAAACCAACCAACAAATTTGATCTTCTGCTTCAAGTGTGATCAATTTAAAGGC 1868
Db 276727 TTTAAACCAACCAACAAATTTGATCTTCTGCTTCAAGTGTGATCAATTTAAAGGC 276668
Qy 1869 GATATTCCCAACCAACAAATTTGATCTTCTGCTTCAAGTGTGATCAAGTGTGATCAAGT 1928
Db 276667 GATATTCCCAACCAACAAATTTGATCTTCTGCTTCAAGTGTGATCAAGTGTGATCAAGT 276608
Qy 1929 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1988
Db 276607 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276548
Qy 1989 GATCAACATTTGATCAACCGTACATTTAAAGCTGAAATCTTCCAAATTAAGCGGAAT 2048
Db 276547 GATCAACATTTGATCAACCGTACATTTAAAGCTGAAATCTTCCAAATTAAGCGGAAT 276488
Qy 2049 GCGGTGTTCTGCAATGTTTCTTCAATTTGAAGGAAATTTGGAAGTCAAGCAATTAAT 2108
Db 276487 GCGGTGTTCTGCAATGTTTCTTCAATTTGAAGGAAATTTGGAAGTCAAGCAATTAAT 276428
Qy 2109 AATGCCAATTTGTTGTTGTCGAATCAACAAATACCATTTGCAAGGCTTCAAGTTGG 2168
Db 276427 AATGCCAATTTGTTGTTGTCGAATCAACAAATACCATTTGCAAGGCTTCAAGTTGG 276368
Qy 2169 AAGGATTAACGCTTTCCTTCAAAAGTGAATTTAAACGATCAAAAGTATTAATTTA 2228
Db 276367 AAGGATTAACGCTTTCCTTCAAAAGTGAATTTAAACGATCAAAAGTATTAATTTA 276308
Qy 2229 CCAAAAACAAATCATGCTTATTAATTTAATGATTAATGCAAGCGCAATTAAT 2288
Db 276307 CCAAAAACAAATCATGCTTATTAATTTAATGATTAATGCAAGCGCAATTAAT 276248
Qy 2288 GGTTTAGCAAACTTAATGCAATGCACTTTAACAATTCAGCGCAATTTAATTAATG 2348
Db 276247 GGTTTAGCAAACTTAATGCAATGCACTTTAATGCAATTCAGCGCAATTTAATTAATG 276188
Qy 2349 AACAATGCGACCAAAATGCAATTAATTAATTTCCGCAATTCAGCTGCAAGCGTGTAT 2408

Db 276187 AACAATGCGACCAAAACAGCAATTAATTAATTTCAATTCAGCGCAAAATGCAAGCTAAT 276128
Qy 2409 AATGCAAACTGAAACGTAATGTCATTTAAAGCAATTCAGCTAATTTCTTTAAATTAAC 2468
Db 276127 AATGCAAACTGAAACGTAATGTCATTTAAAGCAATTCAGCTAATTTCTTTAAATTAAC 276068
Qy 2469 AGCAATTTTCCGCAAAATTCAGGAGGACAAAGGCAACAGTACGTTGAAATGCG 2528
Db 276067 AGCAATTTTCCGCAAAATTCAGGAGGACAAAGGCAACAGTACGTTGAAATGCG 276008
Qy 2529 ACTTGACAACTGCTAGGATACTACATTTGCAATTTAAAGCTTAATTAAGTACATC 2588
Db 276007 ACTTGACAACTGCTAGGATACTACATTTGCAATTTAAAGCTTAATTAAGTACATC 275948
Qy 2589 ACCTTAATTTGCTTATGCTTACGCTTACCTTAACATACGCAAGT---CGCGCTTATTA 2645
Db 275947 ACCTTAATTTGCTTATGCTTACGCTTACCTTAACATACGCAAGT---CGCGCTTATTA 275888
Qy 2646 GAGACGGAACCAACGCAATGCGCAGACATGCTTTCAACATTTGCAAGTAAATGCT 2705
Db 275887 GAGACGGAACCAACGCAATGCGCAGACATGCTTTCAACATTTGCAAGTAAATGCT 275828
Qy 2706 AATTTGATGCGCAAGCAATTCATTTACTTCAATTTCTTATTTGCTTAAAGCGAT 2765
Db 275827 AATTTGATGCGCAAGCAATTCATTTACTTCAATTTCTTATTTGCTTAAAGCGAT 275768
Qy 2766 AATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2825
Db 275767 AATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 275708
Qy 2826 AAGAACCCGAAACCCCTTGACATTAATTTGCTTGAAGAAAGATTAATCAACGCTTA 2885
Db 275707 AAGAACCCGTAACCCCTTGACATTAATTTGCTTGAAGAAAGATTAATCAACGCTTA 275648
Qy 2886 TCGATTAATCTTAATTTACTTTGAATTAATTAATTAATTAATTAATTAATTAATTA 2945
Db 275647 TCGATTAATCTTAATTTACTTTGAATTAATTAATTAATTAATTAATTAATTAATTA 275588
Qy 2946 AATTTGATGAGATTAATTTGCGCAATTTGCTTGAATTAATTAATTAATTAATTAATTA 3005
Db 275587 AATTTGATGAGATTAATTTGCGCAATTTGCTTGAATTAATTAATTAATTAATTAATTA 275528
Qy 3006 CAAATGATTTAGTAAGCAAGCAAGCAAGCAAGCAATTTGAAGCCAAACGATTA 3065
Db 275527 CAAATGATTTAGTAAGCAAGCAAGCAAGCAAGCAATTTGAAGCCAAACGATTA 275468
Qy 3066 CCGATGCTTAAACCAACAAAGTGAAGCCAAATTTGCTTGAAGCCAAACGATTA 3125
Db 275467 CCGATGCTTAAACCAACAAAGTGAAGCCAAATTTGCTTGAAGCCAAACGATTA 275417
Qy 3126 GCGTTTCTGATTAACCGCTGCTGATCAAGCTTGAAGCCATTTGAAGCCAAACGAT 3185
Db 275416 GCGTTTCTGATTAACCGCTGCTGATCAAGCTTGAAGCCATTTGAAGCCAAACGAT 275357
Qy 3186 GAACTGATGCTGAACCAACAAAGTGAAGCCAAATTTGAAGCCAAACGAT 3245
Db 275356 GAACTGATGCTGAACCAACAAAGTGAAGCCAAATTTGAAGCCAAACGAT 275297
Qy 3246 -----GAGTGTCTGATCCCTGCTGATCAAGCTTGAAGCCATTTGAAGCCAAACGAT 3293
Db 275296 GAGTGTGATGCAAGGCTTCTGATTAACCGCTTGAAGCCATTTGAAGCCAAACGAT 275237
Qy 3294 GCGGCACTTGAAGTATTAATGAGCCCAACGATGGAAGAAAGATGCTTGAAGCCAAACGAT 3353
Db 275236 GCGGCACTTGAAGTATTAATGAGCCCAACGATGGAAGAAAGATGCTTGAAGCCAAACGAT 275177
Qy 3354 GAGGCGAAGAAACCAACCAACAAAGTGAAGCCATTTGAAGCCATTTGAAGCCATTTGA 3413
Db 275176 G-----AACAAGGCAACCAACCAACAAAGTGAAGCCATTTGAAGCCATTTGAAGCCATTTGA 275126
Qy 3414 TCGATTAATTCGCAAGATTAATTTGATGCTTCTGCTTGAAGCCATTTGAAGCCATTTGA 3473
Db 275125 TCGATTAATTCGCAAGATTAATTTGATGCTTCTGCTTGAAGCCATTTGAAGCCATTTGA 275066

Db 278167 TATCAAAATGTAAGAAATATATATCAAGCTTGGAGAGAAAGATCTTATATATGCA 278108
Qy 477 GATTACCAATATCCAGATTAACATTAATTTGGTACAGAAAGCGGCTCCAAATGATATGCT 536
Db 278107 GATTATCATATGCTGTTTACATTAATTTGTAATGTAACCTGTAACCTGGGTATGCA 278048
Qy 537 TCGAATATGATGAGAGTACTTATTCAGATAGACAAATATCCAGAACGTGTGTATC 596
Db 278047 ACAAATATGATGAGAAATATATGCTGATAGAGAAATCTATCTGAGCTGTATAT 277988
Qy 597 GGGCTGACGAGAGTTTGGCAATGATCAAGCAAGGCG----- 639
Db 277987 GGGCTGACGAGAGTCAATATGAGGTATAGATTAAGATGAGAAACGAATGATAGTCA 277928
Qy 640 AACAAGTGGCGGTGATATCTATCTCAAGCTGCAATACACATCAAGCTGGA 698
Db 277927 TATATGCTCAAGGTGATATGATATCTTACAGAGAAATCCCATCTCAGAGTGA 277868
Qy 699 GAGGTATATGATATCTGATTTGGAGAGGATGTTGTAAGCGGAGATATATGCTCA 758
Db 277867 AATGATATGATATCAATCAATCTTATGATATGATATGATATGATATGATATGAT 277808
Qy 759 TTAACGATGAGCTCAAGAGGAGAGAGGATGTTCCGATGTTTATATGATGCTGA 818
Db 277807 TTAACGAGGAGGTTCTAAAGGAGATGAGGATGCTGCAATGTTTATATGATGCAAG 277748
Qy 819 AACAAGATGATATTAATGAGATATGAGGAGAGGATGCTTGAAGCAAGAA 878
Db 277747 AAGAAACATATGCTTAAATGCTGATATCAAACTGGGCAATCTTTTGGAGAGCT 277688
Qy 879 AATGATATGATATGATGATGATATCTTATTTGAT-----GAAATTTTC 923
Db 277687 AATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 277628
Qy 924 GAAAGATATTAATATCAATCACTTACACCGAGCTGATATGAGATGATACAAATGAT 983
Db 277627 CCAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 277568
Qy 984 GAAATATGATATGATGATGATATGATATGATATGATATGATATGATATGATATGAT 1043
Db 277567 AATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 277508
Qy 1044 AATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1094
Db 277507 TCAAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 277448
Qy 1095 AATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1154
Db 277447 GTTAAAGCAGACAGGCTATATATTTTACCAACCAAGATGATATGATATGATATG 277388
Qy 1155 TATTTTATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1214
Db 277387 TACCTGGGCAACAGAAAGAAAGAACTTATCAATGAGAAATATATATATCAAGGCT 277328
Qy 1215 GGTGATCTTATTTTGGGATATTTTACATATCTCCAA---ATTCTAACCAATCTGG 1271
Db 277327 GGTGATCTTATTTTGGGATATTTTGTGTAAGGCAAGAAATATATATATCTGG 277268
Qy 1272 CAAGAGCTGCAATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1331
Db 277267 CAAGGTCAGGCTATCTATATGATATGATATGATATGATATGATATGATATGATATG 277208
Qy 1332 GAAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1391
Db 277207 GAAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 277148
Qy 1392 AATTAAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1451
Db 277147 AATTAAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 277088
Qy 1452 CAAGGCAACAAAGCTTATGATATGATATGATATGATATGATATGATATGATATGATATG 1511

Db 277087 GCGGATCAAAAAAGGCTTCAAAAGATGATGATGATGATGATGATGATGATGATGAT 277028
Qy 1512 TTTAAAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1571
Db 277027 TTTAAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 276968
Qy 1572 TTAGATCTTAAAGGAGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1631
Db 276967 TTAGATCTTAAAGGAGATATGATATGATATGATATGATATGATATGATATGATATGAT 276908
Qy 1632 ATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1691
Db 276907 ATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 276848
Qy 1692 AT---TGTATCTATATGATATGATATGATATGATATGATATGATATGATATGATATG 1748
Db 276847 ATATGCTTCAATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 276788
Qy 1749 TACAGGTTGGTTGGCGAAACATATAAATATAACATAGGCGATTAACCTTAT 1808
Db 276787 TACAGGTTGGTTGGCGAAACATATAAATATAACATAGGCGATTAACCTTAT 276728
Qy 1809 TATTAACCAACCAAGATGATATGATATGATATGATATGATATGATATGATATGATATG 1868
Db 276727 TATTAACCAACCAAGATGATATGATATGATATGATATGATATGATATGATATGATATG 276668
Qy 1869 GATATTAACCAACCAAGATGATATGATATGATATGATATGATATGATATGATATGATATG 1928
Db 276667 GATATTAACCAACCAAGATGATATGATATGATATGATATGATATGATATGATATGAT 276608
Qy 1929 AATCATTTAAATTAACCTTGTGATATGATATGATATGATATGATATGATATGATATGAT 1988
Db 276607 AATCATTTAAATTAACCTTGTGATATGATATGATATGATATGATATGATATGATATG 276548
Qy 1989 GATCATGATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 2048
Db 276547 GATCATGATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 276488
Qy 2049 GCGGTGATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 2108
Db 276487 GCGGTGATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 276428
Qy 2109 AATGCAATTTGATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 2168
Db 276427 AATGCAATTTGATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 276368
Qy 2169 ACGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 2228
Db 276367 ACGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 276308
Qy 2229 CCAAAACCAACCAATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 2288
Db 276307 CCAAAACCAACCAATGATATGATATGATATGATATGATATGATATGATATGATATGAT 276248
Qy 2289 GGTATGCAAACTTAAATGATATGATATGATATGATATGATATGATATGATATGATATG 2348
Db 276247 GGTATGCAAACTTAAATGATATGATATGATATGATATGATATGATATGATATGATATG 276188
Qy 2349 ACAAATGCAACCAATGATATGATATGATATGATATGATATGATATGATATGATATGAT 2408
Db 276187 ACAAATGCAACCAATGATATGATATGATATGATATGATATGATATGATATGATATG 276128
Qy 2409 AATGCAAACTTAAATGATATGATATGATATGATATGATATGATATGATATGATATGAT 2468
Db 276127 AATGCAAACTTAAATGATATGATATGATATGATATGATATGATATGATATGATATG 276068
Qy 2469 AGCATATTTGCAACCAATTTAGGAGCAAAAGCAAAAGTACGATGATGATGATGATGATG 2528
Db 276067 AGCATATTTGCAACCAATTTAGGAGCAAAAGCAAAAGTACGATGATGATGATGATGATG 276008
Qy 2529 ACTTGAACCAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 2588
Db 276007 ACTTGAACCAATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 275948

OY	2589	ACGTTAAATACGCTTATTACGACTACTAAACAATACGCAACGCT---CGCGCTCATTA	2645
DB	275947	ACGTTAAATACGCTTATTACGACTACTAAATATACGCAACGCTACCGCGCTCATTA	275888
OY	2646	GAGACGGAAACAACGCAACATCGGAGAACATCGTTCAACATTCAGATGAATGCT	2705
DB	275887	GAGACGGAAACAACGCAACATCGGAGAACATCGTTCAACATTCAGATGAATGCT	275828
OY	2706	AAATTGATGGGAGAGACACTTCACATTAATCACTTATATTGGTATATAAGCAT	2765
DB	275827	AAATTGAGCGGAGAGACACTTCACATTAATCACTTATATTGGTATATAAGCAT	275768
OY	2766	AAATTAATAATATCCATATAGCGTGAAGGCGCTACATATATCTGTCCACACAGGC	2825
DB	275767	AAATTAATAATATCCATATAGCGTGAAGGCGCTACATATATCTGTCCACACAGGC	275708
OY	2826	AAAGAACCCGAAACCTTTGAGCAATTAACTTGGTTGAAGCAAAAGATATCAACCGTTA	2885
DB	275707	AAAGAACCCGTAACCTTTGAGCAATTAACTTGGTTGAAGCAAAAGATATCAACCGTTA	275648
OY	2886	TCAGATAGAGCTCAATTTACTTTAGAAAAATGACAGCGTGATGAGAGGATTAAGCTAT	2945
DB	275647	TCAGATAGAGCTCAATTTACTTTAGAAAAATGACAGCGTGATGAGAGGATTAAGCTAT	275588
OY	2946	AAATTAAGTAAGATGATGCGAATTCGCTTGTGATACCCATTAATAAAGACAGCAATTG	3005
DB	275587	AAATTAAGTAAGATGATGCGAATTCGCTTGTGATACCCATTAATAAAGACAGCAATTG	275528
OY	3006	CACATGATTTAGTATAGAGGAGCAAGGAGAACAACTTTGAAAGCCAAACAGCTTGA	3065
DB	275527	CTCATATATTATGTAAGGAGCAAGGAGAACAACTTTGAAAGCCAAACAGCTTGA	275468
OY	3066	CCGACTGCTTAAACAACAAGGTGAGCCAAAGTGGCTCAAGAAGACAGCGAGACA	3125
DB	275467	CAGACTGCTTAAACAACAAGGTGAGCCAAAGTGGCTCAAGAAGACAGCGAGACA	275417
OY	3126	GCCTTCTGATACCTGCTGCTGATCAAGCCTGTTAAACGATTTGAAGCAACAGCT	3185
DB	275416	GTCTTCTGATACCCCGTCTGCTCAAGCCTGTTAAAGCATTTGAAGCAACAGCT	275357
OY	3186	GAATTAAGCTGTAACAACAAAAGATAGGCAAAAACAAAAGAGCGCTCAAAAAGA	3245
DB	275356	GAAAGACTTCTGTAACAACAAAGATAGGCAAAAACAAAAGAGCGCTCAAAAAGA	275297
OY	3246	-----GAGGTGTTTCTGATCCCTCGCTTGAATCAAGCTGTT---CGCATTAAGA	3293
DB	275296	GCATTAGTGAAGCGCTTTTCTGATACCCCGTGTATCTAAGCAGTTAAAGATTTGCA	275237
OY	3294	GGCGACATTGAGGTTATTTAGTGCCTCCACAGAAATCGGAAAAAGATCGTCAAGCTCAAGA	3353
DB	275233	GTCAAACTTGAAGTTATTAATGCCAACCGAATGTAAAAAGAACTCAAGATCAAG	275177
OY	3354	GAAACGGAAAAACAACGAAACAAAAGACTTGAACGCGTATTAATTAATGTCGTTA	3413
DB	275176	G-----AACAGGCAACAAAAGAAATTGATGACGCGTATCTCAATTAATGTCGTTA	275126
OY	3414	TCAGATATATCTGCAACAGTAATAGTATCTTTCTGTTCAAGATGAATTAATGTCGTT	3473
DB	275125	TCGAGAGTGTCTGCAACAGTAATAGTATGTTTTCGTTCAAGATGAATTAATGTCGTT	275066
OY	3474	TTTGTGATCAAGCAACATCTGCGGTGGTGAACAAATTTGCGACAGATTAATAAGACGCTAT	3533
DB	275065	TTTGTGATCAAGCAACATCTGCGGTGGTGAACAAATTTGCGACAGATTAATAAGACGCTAT	275006
OY	3534	GATTCTGATGCGTTCGCGCTTATCAAGACAGCAAAAGCACTTAACGTCAAAATTGGGGTG	3593
DB	275005	GATTCTGATGCGTTCGCGCTTAT---CAGACAGAAAACGAACTTGCTCAAAATTGGGGTG	274949
OY	3594	CAAAAAGCCTTACATATGACAGAAATTTGGGCGAGTTTCTGACATACCGCTCAATAT	3653
DB	274948	CAAAAAGCCTTACATATGACAGAAATTTGGGCGAGTTTCTGACATACCGCTCAATAT	274889

QY	3654	ACCTTTGATTAACAGATTTAAAAATTCACGGGACATTAAACATGATCGGGTTTGCCAA	3713
Db	274889	ACCTTTGATTAACAGATTTAAAAATTCACGGGACATTAAACATGATCGGGTTTGCCAA	274822
QY	3714	TATCAATGCGGCGATTTTCAATTTGGTGTAAACGTGGGAAACGGGATTCAGTGCAGTAA	3773
Db	274828	TATCAATGCGGCGATTTTCAATTTGGTGTAAACGTGGGAAATTAAGTGCAGTAA	274769
QY	3774	ATGGCTGAGAAACAAGCGCCAAAATTCATCGAAAGCAATAATATAGCGGTAAATGCA	3833
Db	274766	ATGGCTGAGAAACAAGCGCCAAAATTCATCGAAAGCAATAATATAGAGTAATGCA	274705
QY	3834	AGTATTCAGTCCGTTTATAGGCGAATTTGGCAATTCAGCTTAACTTTTGGAGTTAATCGCTAT	3893
Db	274708	AGTATTCAGTCCGTTTATAGGCGAATTTGGGTATTCAGCTTATTTTGGGTATTAATCAT	274644
QY	3894	TTTATGTAACGTGAAAAATATATCATTTGAGAGAGTGAAGTGAAGAGCGCTAGCCTTGCA	3953
Db	274648	TTTATGTAACGTGAAAAATATATCATTTGAGAGAGTGAAGTGAAGAGCGAGCTTGCA	274589
QY	3954	TTTATGCGATTAATAGCTGGGCAATTCAGATGATTAATCATTTACTCCGACAGATTAATTC	4013
Db	274588	TTTATGCGATTAATAGCTGGGCAATTCAGATGATTTATATACCTCCGACAAATTAATTC	274522
QY	4014	AGCGTTAAGCTTATTTCTTGCGAATTTATGTTGATGTTTCAACGCTTAAGTAAACCC	4073
Db	274528	AGCGTTAAGCTTATTTCTTGCGAATTTATGTTGATGTTTCAACGCTTAAGTAAACCC	274469
QY	4074	ACGGTAATTCAGGAGTGTGGCAACACATTTGGAGGTATTTGGCAAAAAAGAGTGGCA	4133
Db	274468	ACGTATTAATGACAGATGTTGCAACAATCTTTGGGCGTTATTTGGCAAAAAAGAGTGGCA	274405
QY	4134	TTAAGGCAAGAAATTTTACATTTCCAAATTTTCGCTTTATCTCAAAATCTCAAGGTTCA	4193
Db	274408	TTAAGGCAAGAAATTTTACATTTCCAAATTTTCGCTTTATTTCAATCTCAAGGTTG	274345
QY	4194	CAACTCGGCAAAACGCAAAATGTGGCGGTGAATTTGGGCTATTCGTGTAAAAATCAACA	4253
Db	274348	CAACTCGGCAAAACGCAAAATGTGGCGGTGAATTTGGGCTATTCGTGTAAAAATCAACA	274289
QY	4254	TAAATTATTCGTTTATTTGATTAACAAGGTGGGTCAATTCAGATTCACACCTTTTATTC	4313
Db	274288	TAAATTATTCGTTTATTTGATTAACAAGGTGGG-----GCAATTCACACCTTTTATTC	274233
QY	4314	AAATAT 4319	
Db	274233	AAATAT 274228	
RESULT 8			
US-1D-080-505-12			
Sequence 12, Application US/1008305			
Patent No. 6676948			
GENERAL INFORMATION:			
APPLICANT: St. Geme, Joseph M.			
TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS			
FILE REFERENCE: A-59941-1/RET/CP/DBR			
CURRENT FILING DATE: 2002-02-22			
PRIOR APPLICATION NUMBER: US 08/296,791			
PRIOR FILING DATE: 1994-10-25			
PRIOR APPLICATION NUMBER: US 09/839,996			
PRIOR FILING DATE: 2001-04-20			
NUMBER OF SEQ ID NOS: 58			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 12			
LENGTH: 5245			
TYPE: DNA			
ORGANISM: Haemophilus influenzae			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (430)..(4740)			
OTHER INFORMATION:			

US-10-080-505-12

Query Match 69.2%; Score 2990.4; DB 4; Length 5245;
 Best Local Similarity 82.0%; Pred. No. 0;
 Matches 3660; Conservative 0; Mismatches 636; Indels 168; Gaps 12;

1 TCATAAGTGGTTAACTA-GTATTTTAAACGAAAAATTACTTAATTAATTAACATT 59
 Db TCATAAGTGGTTAACTAAGTATTTTAAACGAAAAATTACTTAATTAATTAACATT 429
 QY 60 ATGAAAAAACTGATATTCCTGCTTAATTTTAAACGCTGCTCAATTTCAATAGGA 119
 Db ATGAAAAAACTGATATTCCTGCTTAATTTTAAACGCTGCTCAATTTCAATAGGA 489
 QY 120 TCGCAAGCTGGGCTGGTCA-CACTATTTTGGGATGATTAACCAATTAATTCGATATT 179
 Db TCGCAAGCTGGGCTGGTCA-CACTATTTTGGGATGATTAACCAATTAATTCGATATT 549
 QY 180 GCCGAGATTAAGGAGGATTCACAGTTGGGCTCAAAAATTAAAGTTTAAACAAACA 239
 Db GCCGAGATTAAGGAGGATTCACAGTTGGGCTCAAAAATTAAAGTTTAAACAAAG 609
 QY 240 GGGCAATTAGTGGGCAATCAATGACAAAGCCCGATGATTTTCTGTAAGTCA 239
 Db GGGCAATTAGTGGGCAATCAATGACAAAGCCCGATGATTTTCTGTAAGTCA 669
 QY 300 GGTAAAGGCTGGGCACTGGTGAATTAATTAATTTGACGCTGGCAATACGTA 359
 Db GGTAAAGGCTGGGCACTGGTGAATTAATTAATTTGACGCTGGCAATACGCT 729
 QY 360 GGAATTAACAGATTTGATTTTGGGCAAGGAAACAACCCGATCAATGTTTACT 419
 Db GGAATTAACAGATTTGATTTTGGGCAAGGAAACAACCCGATCAATGTTTACT 789
 QY 420 TTAAGATGTTAAAGAAATTAATTAACAAAGAAATTAATTAATTAAGACAT 479
 Db TTAAGATGTTAAAGAAATTAATTAACAAAGAAATTAATTAATTAAGACAT 849
 QY 480 TACCAATTAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 539
 Db TACCAATTAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 909
 QY 540 AATATGAATGCAATCTTATTCAGATTAAGAAACAAATTAATTAATTAATTAATTA 599
 Db AATATGAATGCAATCTTATTCAGATTAAGAAACAAATTAATTAATTAATTAATTA 969
 QY 600 TCTGAGGAGGAGTTTGGGAAATGATCAAGCAAGGAGCAAA----- 644
 Db TCTGAGGAGGAGTTTGGGAAATGATCAAGCAAGGAGCAAA----- 1029
 QY 645 GTTGCAGTGCATATCATTAATTCAGACGCTGGCAATCAACATCAAGCTGAGACAT 704
 Db GTTGCAGTGCATATCATTAATTCAGACGCTGGCAATCAACATCAAGCTGAGACAT 1089
 QY 705 AATGAATTCGATTTGGGAGGCGATTTGTAAGCGGGAGATTTGCTCAATACG 764
 Db AATGAATTCGATTTGGGAGGCGATTTGTAAGCGGGAGATTTGCTCAATACG 1149
 QY 765 AATGAGGCTCAAGGAGGAGGAGGCTTCGAGATTTTAAATTAATTAATTAATTA 824
 Db AATGAGGCTTCGTTGGAGATTTGCTTCGAGATTTTAAATTAATTAATTAATTA 1209
 QY 825 AATGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 884
 Db AATGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1269
 QY 885 TTTCAATTTGTTGCAATCTTATTTTGAAGAA---TTTTCGAAAGATTTAATCA 941
 Db TTTCAATTTGTTGCAATCTTATTTTGAAGAA---TTTTCGAAAGATTTAATCA 1329
 QY 942 TCACTTAACCGGAGCTGTAATGAGTGAACATTAATTAATTAATTAATTAATTA 1001
 Db TCACTTAACCGGAGCTGTAATGAGTGAACATTAATTAATTAATTAATTAATTA 1389

QY 1002 GGGTCTAATAC------AGAAATCAGAAATCCATCAGAAATTAATTAATTA 1052
 Db GGGTCTAATAC------AGAAATCAGAAATCCATCAGAAATTAATTAATTAATTA 1449
 QY 1053 GCAATATAGCTTACCTTTGAAAGAGAGATTAAGCTTCAATCTCAATATGACGA 1112
 Db GCAATATATAGAGATTAAGAGAGAGATTAAGAGAGATTAAGAGAGATTAAGAGAG 1509
 QY 1113 CCTAATATTTCTCCAGCTTTAAACAAATGAGAAACGCTAATTAATTAATTAATTA 1172
 Db GTAATATCTTAAACCAACATTAATTAATTAATTAATTAATTAATTAATTAATTA 1569
 QY 1173 CAAGATCATTAACCTTGATGATCAATTAACCAAGGCGGAGTCTTTATTTGAG 1232
 Db ACGAGAACTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1629
 QY 1233 GGTAAATTTAAGATTCCTCAATTTCAACCAACTTGGCAAGAGCTGGCATACATGA 1292
 Db GGTAAATTTAAGATTCCTCAATTTCAACCAACTTGGCAAGAGCTGGCATACATGA 1689
 QY 1293 AGTAAATTAAGACGCTTACCTGAAAGTAAATGCGTGGAAACATGATGCACTTTCA 1352
 Db AGTAAATTAAGACGCTTACCTGAAAGTAAATGCGTGGAAACATGATGCACTTTCA 1749
 QY 1353 ATTGTTAAAGAGATTCAGCTTCAAGCAAGGCAAAATTAAGTTCAATCAGCTTA 1412
 Db ATTGTTAAAGAGATTCAGCTTCAAGCAAGGCAAAATTAAGTTCAATCAGCTTA 1809
 QY 1413 GCGATATGTAAGTATTTGAGAGCGAGGAGCAATCAAGCAACAGCTTT 1472
 Db GCGATATGTAAGTATTTGAGAGCGAGGAGCAATCAAGCAACAGCTTT 1869
 QY 1473 AGTAAATTTGCTTTAGAGGAGAGGAGCTGTAATTAAGATTAAGATTAAGATTA 1532
 Db AGTAAATTTGCTTTAGAGGAGAGGAGCTGTAATTAAGATTAAGATTAAGATTA 1929
 QY 1533 GATACCGAATTAATTTTATTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 1592
 Db GATACCGAATTAATTTTATTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 1989
 QY 1593 TTAACCTTTAAAGTATTCGAATTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1652
 Db TTAACCTTTAAAGTATTCGAATTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2049
 QY 1653 ACTCAAGGCGCTAATTTGCTTATTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 1709
 Db ACTCAAGGCGCTAATTTGCTTATTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 2109
 QY 1710 AATTAATTAATTAATTAATTTTAAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 1769
 Db AATTAATTAATTAATTAATTTTAAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 2169
 QY 1770 ACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1829
 Db ACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2229
 QY 1830 GGTACTTTGCTTACCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1889
 Db GGTACTTTGCTTACCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2289
 QY 1890 AATCTATTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1949
 Db AATCTATTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2343
 QY 1950 TCGAAATTAAGAGGATTAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2009
 Db TCGAAATTAAGAGGATTAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2403
 QY 2010 ACATTTAAAGCTGAAGATTCGAAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2069
 Db ACATTTAAAGCTGAAGATTCGAAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2463

2070 TCTCAATTGAGGAAATTGGACAGTCAGCAATTAATGCAAAATGTCACATTTGGTGTG 2129
2464 TCTCAATTGAGGAAATTGGACAGTCAGCAATTAATGCAAAATGTCACATTTGGTGTG 2523
2130 CCAATTCAGCAAAATTCATTTGACAGCGTTGAGTGGACAGGATTAACGACTTGTCA 2189
2524 CCAATTCAGCAAAATTCATTTGACAGCGTTGAGTGGACAGGATTAACGACTTGTAA 2583
2190 AAGTGGATTTAACGGATTCAGCAAAATGTTATTAATTTCAATCAAAAACAAATCAATG 2249
2584 ACTGGATTTAACGGATTCAGCAAAATGTTATTAATTTCAATCAAAAACAAATTAATG 2643
2250 TCTATTAATTATCTGATTAATGCAAGCGCAATGTTAAAGTTAGCAAACTTAATG 2309
2644 TCTATTAATTATCTGATTAATGCAAGCGCAATGTTAAAGTTAGCAAACTTAATG 2703
2310 AATGTCATTTAAACAAATTCAGCAATTTAATTAATTAATGCAATGCAACCAATAG 2369
2704 AATGTCATTTAAATTCATGCAATTTAATTAATTAATGCAATGCAACCAATAG 2763
2370 AATATGCACTTTCCGCAATTTCAATGCAAGCGTGGATTAATGCAAACTTTAGCGT 2429
2764 AATATGCACTTTCCGCAATTTCAATGCAAGCGTGGATTAATGCAAACTTTAGCGT 2823
2430 GTGCATTTAAGGATTCAGTCATTTCTTTAAAAACAGCATTTTGGACCAAT 2489
2824 GTGCATTTAAGGATTCAGTCATTTCTTTAAAAACAGCATTTTGGACCAAT 2883
2490 CAGGAGACAAAGCAACAGTCAGCTTGGAAAATGCGACTTGGCAATTCCTAGCAT 2549
2884 CAGGAGACAAAGCAACAGTCAGCTTGGAAAATGCGACTTGGCAATTCCTAGCAT 2943
2550 ACTGATTTGAGATTTTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 2609
2944 GCGCATTTGAGATTTTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 3003
2610 GCTAGCTCAAAACATTCAGCAAGT---CGCGCTTATGAGACGCAAAACAGCGCA 2666
3004 GCTAGCTCAAAACATTCAGCAAGT---CGCGCTTATGAGACGCAAAACAGCGCA 3063
2667 TCGCAGACATTCGTTCAACATTCAGCAAGT---CGCGCTTATGAGACGCAAAAC 2726
3064 TCGCAGACATTCGTTCAACATTCAGCAAGT---CGCGCTTATGAGACGCAAAAC 3123
2727 TTTCAATTTACTTCAATTTTGTGCTAATTAATTAATTAATTAATTAATTAATTA 2786
3124 TTTCAATTTACTTCAATTTTGTGCTAATTAATTAATTAATTAATTAATTAATTA 3183
2787 GCTGAGGCGATTAATTAATTCGTTGCAACAGCAAGCAAGCAAGCAAGCAAGCA 2846
3184 GCTGAGGCGATTAATTAATTCGTTGCAACAGCAAGCAAGCAAGCAAGCAAGCA 3243
2847 CAATTTAATTTGTTGAAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTA 2906
3244 CAATTTAATTTGTTGAAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTA 3303
2907 TTAGAAATGCAAGTGGATTCAGCAAGTGGATTAATTAATTAATTAATTAATTAAT 2966
3304 TTAGAAATGCAAGTGGATTCAGCAAGTGGATTAATTAATTAATTAATTAATTAAT 3363
2967 GAATTCGCTTGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3026
3364 GAATTCGCTTGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3423
3027 GAGCAAGCAAGCAAGTGGATTCAGCAAGTGGATTAATTAATTAATTAATTAATTA 3086
3424 GAGCAAGCAAGCAAGTGGATTCAGCAAGTGGATTAATTAATTAATTAATTAATTA 3483
3087 GGTGAGCAAGCAAGTGGATTCAGCAAGTGGATTCAGCAAGTGGATTCAGCAAG 3119
3484 AGTATGCAAGCAAGTGGATTCAGCAAGTGGATTCAGCAAGTGGATTCAGCAAG 3543
3120 ----- 3119

3544 CAGTTAGAGATTTACAAAGCCGAACAGTTGACCGACTGCTGAAAAACAAAAATTAAG 3603
3120 -----NAGCAGCGTTTCTGATTAATTAATTAATTAATTAATTAATTAATTA 3158
3604 GCAAAAAAGTGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3663
3159 TTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3218
3664 TTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3723
3219 AAAACAAAAAGTGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3278
3724 AAAAAGCGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3783
3279 CTGTC---GATTAAGAGCGGATTAATTAATTAATTAATTAATTAATTAATTAAT 3335
3784 CCGTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3843
3336 GATGCTAGCTCAAGAAAGCGGAAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3395
3844 GAACTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3894
3396 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3455
3895 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3954
3456 GATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3515
3955 GATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4014
3516 CAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3575
4015 CAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4071
4072 TTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4131
3636 CATAGCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3695
4132 CATAGCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4191
3696 ATGTCGAGTTTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3755
4192 ATGTCGAGTTTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4251
3756 GAAATCAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3815
4252 GAAATCAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4311
3816 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3875
4312 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4371
3876 TTTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3935
4372 TTTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4431
3936 AAAAGCTTACGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3995
4432 AAAAGCTTACGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4491
3996 ACTTCAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4055
4492 ACTTCAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4551
4056 AAGCTTAAGTAAACCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4115
4552 AAGCTTAAGTAAACCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4611
4116 TGGCAAAAAGTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4175

Db 4612 TGGCAAAAGAGGAGATTAAGCGAAATTTTACATTTCCAACTTCCTGCTTTAT 4671
 Qy 4176 TCAAAATCTCAAGTTCAACATCGGCAACAGCAAAATGCGGCGTGAATTTGGGCTAT 4235
 Db 4672 TCAAAATCTCAAGTTCAACATCGGCAACAGCAAAATGCGGCGTGAATTTGGGCTAT 4731
 Qy 4236 CTTGGTAAATCAACATATTTATGTTATTTGATTAACAAGTGGGTGAGATCGA 4295
 Db 4732 CTTGGTAAATCAACATATTTATGTTATTTGATTAACAAGTGGGTGAGATCGA 4788
 Qy 4296 TCCCACTTTTATTTCCAAATAT 4319
 Db 4789 --CCACCTTTTATTTCAATAT 4810

RESULT 9

US-10-080-505-16
 ; Sequence 16, Application US/10280505
 ; Patent No. 6676948
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/FRT/DCP/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 4828
 ; TYPE: DNA
 ; ORGANISM: Haemophilus influenzae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (313)..(4548)
 ; OTHER INFORMATION:
 US-10-080-505-16

Query Match 68.8%; Score 2970; DB 4; Length 4828;
 Best local Similarity 82.3%; Pred. No. 0;
 Matches 3581; Conservative 0; Mismatches 645; Indels 124; Gaps 10;

Qy 1 TCAATAGCTTTAATCTA-GTATTTTAAAGAAAATTTACTTATTAATAACAT 59
 Db 253 TCAATAGCTTTAATCTA-GTATTTTAAAGAAAATTTACTTATTAATAACAT 312
 Qy 60 ATGAAAAAACTGATTTGCTTAAATTTTAAACGCTTCATTTCAATTAAGGATGTA 113
 Db 313 ATGAAAAAACTGATTTGCTTAAATTTTAAACGCTTCATTTCAATTAAGGATGTA 372
 Qy 120 TCGCAAGCTGGGCTGTCACATTAATTTTGGATTTTACCAATTAATGCGATTTT 179
 Db 373 TCGCAAGCTGGGCTGTCACATTAATTTTGGATTTTACCAATTAATGCGATTTT 432
 Qy 180 GCCGGAATTAAGGGAAGTTCAAGTTGGGGCTCAAAATTAAGTTTAAACAATA 239
 Db 433 GCCGGAATTAAGGGAAGTTCAAGTTGGGGCTCAAAATTAAGTTTAAACAATA 492
 Qy 240 GGGCAATTAAGTTGGCAATCATATGCAAAAGCCCGATGATTTTCTGATGTC 299
 Db 493 GGAATTTAAGTTGGCAATCATATGCAAAAGCCCGATGATTTTCTGATGTC 552
 Qy 300 GCTAAGCGCTGGCAAGCTTTGGTGAATTAATATATTTAGCGTGGCAATACGTA 359
 Db 553 CCAATAGCGCTGGCAAGCTTTGGTGAATTAATATATTTAGCGTGGCAATACGTA 612
 Qy 360 GGAATTAAGTTGATTTTGGTGAAGGGAACAACCCGATCAAGCTGTTTACT 419
 Db 613 GGAATTAAGTTGATTTTGGTGAAGGGAACAACCCGATCAAGCTGTTTACT 672

Qy 420 TATAGATTGTTAAACGAATTAATCAAAAAAGATATTTACTCTTATGAGACAT 479
 Db 673 TATAGATTGTTAAACGAATTAATTAATAACATTAAGCTATCTTATGAGAAAGAC 732
 Qy 480 TACCATTAATCAAGATTAATTAATTTGTTACAGAGCGCTTCAATGATATGACTTCG 539
 Db 733 TACCATTAATCAAGCTTACATTAATTTGTTACAGAGCGCTTCAATGATATGACTTCG 792
 Qy 540 AATATGATGAGTACTTTTACAGTAGAACAAATTTTCAGAACCGTTCGATCGC 599
 Db 793 GATATGATGAGTACTTTTACAGTAGAACAAATTTTCAGAACCGTTCGATCGC 852
 Qy 600 TCTGACAGCGATTTTGGCGAATGATCAAGCAAAAGCGCAACAAAGTTCCGCGTATAT 659
 Db 853 TCCGCGTGGAGTTTGGCGAATGATCAAGCAAAAGCGCAACAAAGTTCCGCGTATAT 912
 Qy 660 CATTAATCTGACAGCTGCAATACACAAATGAGGTGAGCAAGTATGATTTGAT 719
 Db 913 CATTAATCTGACAGCTGCAATACACAAATGAGGTGAGCAAGTATGATTTGAT 972
 Qy 720 TTGGAGCGGATTTGTTAAAGGGAGAAATATGTCATTTCCGATTCGAGCTCAAG 779
 Db 973 CTGAGCGGATTTGTTAAAGGGAGAAATATGTCATTTCCGATTCGAGCTCAAG 1032
 Qy 780 GGGGACAGTGTCTCCGATTTTATGATCTGAAATAAATAATGTTATTAAT 839
 Db 1033 GGGGACAGTGTCTCCGATTTTATGATCTGAAATAAATAATGTTATTAAT 1092
 Qy 840 GGGATTTACGGGAAGCAACCTTTTGAAGCAAGAAATGAGTTCAATGGTTCC 899
 Db 1093 GGGATTTACGGGAAGCAACCTTTTGAAGCAAGAAATGAGTTCAATGGTTCC 1152
 Qy 900 AATCTTATTTGATGAAATTTTCGAAAGATTTAATCAATCACTTACACCAAGCT 959
 Db 1153 AATCTTATTTGATGAAATTTTCGAAAGATTTAATCAATCACTTACACCAAGCT 1212
 Qy 960 GGTATGAGTGTACCAATTAATGGAATGATATGAGGCTTATTAATCAAGAA 1019
 Db 1213 GGTATGAGTGTACCAATTAATGGAATGATATGAGGCTTATTAATCAAGAA 1272
 Qy 1020 TCAAGGATTCATCAAGAAATTAATGTTAGCAAAATGAGTTTACCTTGAAGAG 1079
 Db 1273 TCAAGGATTCATCAAGGATTCATGAGTTTGAAGCAAAATGAGTTTACCTTGAAGAG 1332
 Qy 1080 AAGGATTAAGTCAATTCCTAGATTAAGAGGCTTAATTTTCCGAGCTTAAAC 1139
 Db 1333 AAGGATTAAGTCAATTCCTAGATTAAGAGGCTTAATTTTCCGAGCTTAAAC 1386
 Qy 1140 AATGAGAAAGCTATATTTTATGATCAAAACAAAGATCAATTAATCTGCAATGAC 1199
 Db 1387 AATGAGAAAGCTATATTTTATGATCAAAACAAAGATCAATTAATCTGCAATGAC 1446
 Qy 1200 ATTACCAAGGGGCGGTGCTTTATTTTGAAGGTAATTTTACAGTATCTCAAAATCT 1259
 Db 1447 ATTACCAAGGGGCGGTGCTTTATTTTGAAGGTAATTTTACAGTATCTCAAAATCT 1506
 Qy 1240 AACCAAACTGGCAAGAGGCTGGCATACATGTAAGTAATAAGCAAGCTTACTTGA 1319
 Db 1507 AACCAAACTGGCAAGAGGCTGGCATACATGTAAGTAATAAGCAAGCTTACTTGA 1566
 Qy 1320 GTAAATGCGTGAACATGATCGATTTCTAAATTTGTTAAAGGAACAATTCAGCTTCA 1379
 Db 1567 GTAAATGCGTGAACATGATCGATTTCTAAATTTGTTAAAGGAACAATTCAGCTTCA 1626
 Qy 1380 GCTAAAGGGGAATTAAGTTGATCAAGCTGAGCGATGTAAGTCAATTTTGAAGCAG 1439
 Db 1627 GCTAAAGGGGAATTAAGTTGATCAAGCTGAGCGATGTAAGTCAATTTTGAAGCAG 1686
 Qy 1440 CAGGACAGATCAAGGCAACAAGCTTTAGTGAATCTGGCTGGTTAGCGGAG 1499
 Db 1687 CAGGACAGATCAAGGCAACAAGCTTTAGTGAATCTGGCTGGTTAGCGGAG 1746
 Qy 1500 GGGAGCTGTCAATTAACGATGATTAACATTTGATGATTAATTTTATTTGAGCTTT 1559

Db	1747	GGTACGGTTACGTTAAACGATGACACACGATTTATATCTGATTAATTTATTTCCGGCTTC	1806
Qy	1560	CGTGTGTGTGCTTAGATCTTTAACGGGCACTTCATTAACCTTTAAACGTATCCAAATACG	1619
Db	1807	CGTGTGTGTGCTTAGATCTTTATATGGGCACTTCATTAACCTTTAAACGTATCCAAATACG	1866
Qy	1620	GACACGGGGGCGCATGATTTGTGACCATTAATACATCAAGCGCGTAAATGTGCACTTTACT	1679
Db	1867	GATACGGAGACCAACGATGTTTATATCACAAATGACCAACAGATCTTACAGTACACTTACT	1926
Qy	1680	GGGAAACGAAACGATTTGTCTACTATATGAGAAATATATTAATTAACCTGATTTACAGAAA	1739
Db	1927	GGCAGCGATACCTTAATATGACAAACCTGGGATTTAAACCAATTAAGTGAT-----	1977
Qy	1740	GAATTTGCTTACACGGTTGGTTGGCGAAACAGATAAAAATTAACACATGGCGGATTA	1789
Db	1978	--ATTGCTTTTATGT	2034
Qy	1800	AACCTTTATTAATAACCAACCAAGAGATGTACTTTGCTACTTTCAAGTGGTACAAAT	1859
Db	2035	ATATGTTACTTACATCCGCTTTACAAAGATATACCTTCCTCATAGGTGGAGCAAT	2094
Qy	1860	TTAAAAAGCGATATTACCCCAACAAAGGTAACTATTTTTCACGCGTACGACCG	1919
Db	2095	TTAAAAAGCGCATATTACTACAGACGCTGTGACCTTAAGTTATGTGTCCGCCAACACCA	2154
Qy	1920	CACGCTTACATATTAAATTAACGTTGTGTGAAATATGAAAGGTATATACACAGCGCA	1979
Db	2155	CACGATACATCAATTAA-----TCGCTTAACGAGCTTGGCGACCTTAAGGCGCA	2208
Qy	1980	ATTGTGTGTGATCAAGATTTGATCAACCGTACATTTAAAGCTGAAACTTCCAAATTAA	2039
Db	2209	GTGTGTATTGATACGATTTGATCAACCGTACATTTAAAGCTGAAACTTCCAAATTAA	2268
Qy	2040	GGCGGAATGTGCGTGTGTCTCTGTGCAATGTTCTTCAATTGAGGAAATTTGGACATGTACG	2099
Db	2269	GGCGGAATGTGCGTGTGTCTCTGTGCAATGTTCTTCAATTGAGGAAATTTGGACATGTACG	2328
Qy	2100	AATTAATGCAATGTGCACATTTGTGTGTGTGCAATCAACAAATACCATTTGACGCGT	2159
Db	2329	AATTAATGCAATGTGCACATTTGTGTGTGTGCAATCAACAAATACCATTTGACGCGT	2388
Qy	2160	TCAGATTGACAGGATTTAACGCTTGTCAAAAAGTGATTTAACCGATCAAAAGTTAT	2219
Db	2389	TCAGATTGACAGGATTTAACGCTTGTCAAAAAGTGATTTAACCGATCAAAAGTTAT	2448
Qy	2220	AATTTATACCAAAAACAAATCAATGAGCGCTTAATTAATCAATGATATGCAACGCG	2279
Db	2449	GATTCATACCGCAACGCAATTTGTGTGTGTGTGCAATCAACAAATACCATTTGACGCGT	2388
Qy	2280	AATGTTAAAGTTTAGCAAACTTATATGCAATGTCACTTTAACAAATCAACGCAATTT	2339
Db	2509	AATATTCATGTTTAGCAAACTTATATGTAATGTCACTTTAAATCAATACGCAATTT	2568
Qy	2340	AATTTAACCAACATGTCAACCAATATAGGCAATTTGCACTTTCCGCAATTTACATGCA	2359
Db	2569	AATTTAACCAACATGTCAACCAATATAGGCAATTTGCACTTTCCGCAATTTACATGCA	2678
Qy	2400	ACGGTGAATATGCAAACTTGAAACGCTATATGTCACTTTAACGATTCAGTCAATTTCT	2459
Db	2629	ACGGTGAATATGCAAACTTGAAACGCTATATGTCACTTTAACGATTTCTGTCATTTCT	2688
Qy	2460	TTAAAAAACGCACTTTTTCGACCAATTCAGGAGACAAAGGTCACACGATGTG	2519
Db	2689	TTAAAAAACGCACTTTTTCGACCAATTCAGGAGACAAAGGTCACACGATGTG	2748
Qy	2520	GAATATGCAATTTGCAACATGCTGTGCACTACTACATTTGCAATTTAAAGCTAAATAC	2579
Db	2749	GAATATGCAATTTGCAACATGCTGTGCACTACTACATTTGCAATTTAAAGCTAAATAT	2808
Qy	2580	AGTACGATCACTTAAATTCAGCTTATTCAGCTTACCTTACCAATATGCGCACT--GCG	2636

[illegible]

```

QY 3633 TCGCATAGCCGTTGAGATATACCTTTGATGAAACAGTTTAAATACAGGACATTACG 3692
DB 3937 TCGCATAGCCGTTGAGATATACCTTTGATGAAACAGTTTAAATACAGGACATTACG 3996
QY 3693 ATGATGTCGGGTTTTCGCCAATATCAATGAGGCGAATTTTCAATTTGGTGAACGTCGGA 3752
DB 3997 ATGATGTCGGGTTTTCGCCAATATCAATGAGGCGAATTTTCAATTTGGTGAACGTCGGA 4056
QY 3753 ACGGGAATCAGTCCGATTAATATGCTGAAGAACAAACGCAAAATATTCCTGAAAAGCG 3812
DB 4057 ACGGGAATGATGCGATTAATATGCTGAAGAACAAACGCAAAATATTCCTGAAAAGCG 4116
QY 3813 AATAATATGCGCGTAAGTGAAGTTATCACTTCGTTTGAAGGCAATGAGGCAATGAGCT 3872
DB 4117 AATAATATGCGCGTAAGTGAAGTTATCACTTCGTTTGAAGGCAATGAGGCAATGAGCT 4176
QY 3873 TATTTGAGTATATGCTATTTTATTTGAACGTGAATATTCATCTGAGAGTGA 3932
DB 4177 TATTTGAGTATATGCTATTTTATTTGAACGTGAATATTCATCTGAGAGTGA 4236
QY 3933 GTGAAAACGCTTACGCTTTCATTTATGCTATATATGCTGCAATTGAGTTATATCA 3992
DB 4237 GTGAAAACGCTTACGCTTTCATTTATGCTATATATGCTGCAATTGAGTTATATCA 4296
QY 3993 TTTACTCGGACATATATATCAGCGTTACGCTTATTTCTTCTGCAATTATGAGTAT 4052
DB 4297 TTTACTCGGACATATATATCAGCGTTACGCTTATTTCTTCTGCAATTATGAGTAT 4356
QY 4053 TCAAAACGCTATACGTAACAAACGCTTAATCTCAGGTTTGAACCAACATTTGAGCT 4112
DB 4357 TCAAAACGCTATACGTAACAAACGCTTAATCTCAGGTTTGAACCAACATTTGAGCT 4416
QY 4113 TATTGGCAAAAGATGAGATTAAGGCAAAATTTTACCTTTCCAAATTTCCGCTTTT 4172
DB 4417 TATTGGCAAAAGATGAGATTAAGGCAAAATTTTACCTTTCCAAATTTCCGCTTTT 4476
QY 4173 ATCTCAAAATCTCAAGGTTTCACAACTGCGCAACAGCAAAATGAGGCGTGAATTTGGC 4232
DB 4477 ATCTCAAAATCTCAAGGTTTCACAACTGCGCAACAGCAAAATGAGGCGTGAATTTGSA 4536
QY 4233 TATCGTTGTTAAATCAACATTAATTTAT 4262
DB 4537 TATCGTTGTTAAATCAACATTAATTTAT 4566

```

RESULT 10
 US-10-080-505-14
 ; Sequence 14, Application US/10080505
 ; Patent No. 6676948
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Geme, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RFI/DCF/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; PRIORITY FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1.
 ; SEQ ID NO 14
 ; LENGTH: 4822
 ; TYPE: DNA
 ; ORGANISM: Haemophilus influenzae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (388)..(4563)
 ; OTHER INFORMATION:
 US-10-080-505-14

Query Match 64.1%; Score 2770.2; DB 4; Length 4822;

Best Local Similarity 79.6%; Pred. No. 0;
 Matches 3464; Conservative 0; Mismatches 813; Indels 76; Gaps 13;

```

QY 1 TCAATATGCTTTTACCTA-CTATTTTATATGCAAAATATCTTATTAATTAATTAAT 59
DB 328 TCAATATGCTTTTACCTA-CTATTTTATATGCAAAATATCTTATTAATTAATTAAT 387
QY 60 ATGAAAAAATGATTTTGTCTTAATTTTAAACGCTGATTCATTAAGGATATGTA 119
DB 388 ATGAAAAAATGATTTTGTCTTAATTTTAAACGCTGATTCATTAAGGATATGTA 447
QY 120 TCGCAAGCTGGGCTGCTACACTTATTTTGGATGATTAACAAATTAATGATTTT 179
DB 448 TCGCAAGCTGGGCTGCTACACTTATTTTGGATGATTAACAAATTAATGATTTT 507
QY 180 GCGAGATTAAGGAAATTCACAGTTGAGGCTCAAAATTTTACGTTTATTAACAA 239
DB 508 GCGAGATTAAGGAAATTCACAGTTGAGGCTCAAAATTTTACGTTTATTAACAA 567
QY 240 GGGCAATTAATGAGCATCATCATGACAAAGCCCGATGATGATTTTCTGATGCTCA 299
DB 568 GGGCAATTAATGAGCATCATCATGACAAAGCCCGATGATGATTTTCTGATGCTCA 627
QY 300 CGTACGCGCTGAGCAAGCTTGTGTAATATATATATATATATATATATATATATAT 359
DB 628 CGTACGCGCTGAGCAAGCTTGTGTAATATATATATATATATATATATATATATAT 687
QY 360 GATATATACGATTTTATTTTGTGATGAGGAAACACCCGATCAACATGCTTACT 419
DB 688 GATATATACGATTTTATTTTGTGATGAGGAAACACCCGATCAACATGCTTACT 747
QY 420 TATTAATGTTTAAACGAATTAATCAAAATTAATTAATTAATTAATTAATTAATTA 479
DB 748 TATTAATGTTTAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 801
QY 480 TACCATATACACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 539
DB 802 TACCATATACACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 861
QY 540 AATATATACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 599
DB 862 AATATATACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 921
QY 600 TCTGACGAGCAATTTTGGCCAAATGATCAGACAAAGGC---GACCAAGTCCCGGTCA 656
DB 922 AGTGGCATCAATGGTGAAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 981
QY 657 TATCATTAATCAAGCTGCAATTAACCAATCAAGCTGCAATTAATTAATTAATTAAT 716
DB 982 GGTTCATGTTTATAGGTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1041
QY 717 TATTTGGAGGAGGATTTGCTTAAGCGGAGATTAATTAATTAATTAATTAATTAAT 776
DB 1042 GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1101
QY 777 AAGGGGAGCAAGTGTCTCCGATGTTTATTAATTAATTAATTAATTAATTAATTAAT 836
DB 1102 TTGGGGAGATAGGTGTCTCCGATGTTTATTAATTAATTAATTAATTAATTAATTAAT 1161
QY 837 AATGGGATTTTACGGAAGCAACCTTTTGAAGGCAAGAAATGAGTTTCAATTTGCTT 896
DB 1162 AATGGGATTTTACGGAAGCAACCTTTTGAAGGCAAGAAATGAGTTTCAATTTGCTT 1221
QY 897 GCGAATCTTATTTGATGAATTTTTCAGAAAGATTAATTAATTAATTAATTAATTAAT 956
DB 1222 GCGAATCTTATTTGATGAATTTTTCAGAAAGATTAATTAATTAATTAATTAATTAAT 1281
QY 957 GCTGTATATGAGTGTACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1006
DB 1282 AATGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1341
QY 1007 -TATTAATCAAGAAATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1064

```

Db 1342 AGTACATTACTTAAAGGCTATTCAGCTGAGCAATGAGGCTCTTTATGATTAATAGC 1401
 Qy 1065 TTACCTTGAAAGAGATTAAT-----GTTCAATATCTGATATGAGGACTTAAT 1118
 Db 1402 CAATTCATGATGCTAGAGATTAATATGCGATGATATCTCCCTTATTAAGGCTTAAT 1461
 Qy 1119 ATTATCTCCAGGTTTAAACATGAGAAAGCATATATTTATGATCAAAAACAGGA 1178
 Db 1462 CCAATGCGCAGCATTAACATGAGAAAGATTAATCTTGGCAGTACAGGAACAGGA 1521
 Qy 1179 TCATTAATCTTGCACTGACATTAACAGAGGCGGCTGCTTATTTGAGGTAAT 1238
 Db 1522 ACTTAAACATGAAATATATTAATTAATCAAGCTGAGTGAATGATTTGAAGGTAAT 1581
 Qy 1239 TTTACGATATCCCAAT---TCATACCAATCTGGCAGAGAGCTGGCATACGTAAGT 1295
 Db 1582 TTGTTGTAAAGGCAATCAAAATATATATATATCTGGCAAGGTCAGAGCTTTCTGTGGA 1641
 Qy 1296 GAAATAGCAGCGTTACTTGGAAAGTAAATGCGCTGAGACATGATGACTTTCTTAATTT 1355
 Db 1642 GAAGAAAGTACTGTTGAATGCGAGTGCATATTCAGAGGCGATGCTTATCCAAATTT 1701
 Qy 1356 GGTAAAGGAACATGCACTTCAAGCCAAAGGGGAAATTAAGGTTGATCAGGCTAGC 1415
 Db 1702 GGGCTGGAAACCTTACTGTTAATGTTAAAGGAAATTAAGGAAAGCTGAGTGGGT 1761
 Qy 1416 GATGTAAGTCAATTTGAGAGGAGGAGGAGCATCAAGGCAACAAACAGCCTTAAGT 1475
 Db 1762 AACGTTGGTTGTGTGATCAACAGAGATGATCAGTCAAAACAGCCTTTAA 1821
 Qy 1476 GAAATGGCTGGTTAGCGGAGAGGAGCTGTTCAATTAACGATGATTAACAAATTTGAT 1535
 Db 1822 GAAGTTGCAATGTAAGTGTAGAGTACCGTTCACTAAATAGTGCATACAGTTGAT 1881
 Qy 1536 ACCGATTAATTTTATTTGCTGCTTCTGTGTGCTGCTTAATCTTAAGGCTTAATCA 1595
 Db 1882 CTTAAACAAATTTTATTTGCTGCTTCTGTGTGCTGCTTAATCTTAAGGCTTAATCA 1941
 Qy 1596 ACCTTAAAGTATCCAAATATAGGAGGAGGAGGAGATGATGTAACCATTAATCAACT 1655
 Db 1942 ACCTTAAAGTATCCAAATATAGGAGGAGGAGGAGATGATGTAACCATTAATCAACT 2001
 Qy 1656 CAAGCGCTAATGCTACTATTAATGAGAAAGCAATGTTCTTAATGAGAAATTAAT 1715
 Db 2002 CAAGCGCAAAATTAATGATTAACAGGCAAGCACTTAATTAATCA-----GAT 2049
 Qy 1716 ATTAATTAATGATTAACAGAAAGAAATTTGCTTAACAGGTTGTTGGGAAACGAT 1775
 Db 2050 AGCAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2109
 Qy 1776 AAAAATTAACAAATGAGGAGGATTAACCTTATTTAATCAACAGCAAGAGATGCTACT 1835
 Db 2110 AAAAGCTAAACAAATGAGGATTAATGATTAATTAATTAATTAATTAATTAATTAAT 2169
 Qy 1836 TTGCTACTTTCAGGTGTATCAAAATTTAAAGGCGATTAATCCAAACAAAGGTAAC 1895
 Db 2170 TTGTTGCTTTCTGGGGGAGCAATTTAAAGGCAATTAACAGCAAAATGTTGATAC 2229
 Qy 1896 TTTTTCAGCGGTAGACGACACGCGACGCTCAATCAATTTAATTAAGTTGTGAGAA 1955
 Db 2230 GTTTTATGAGTGTGTGCAACGCTCAATGCTTAATTAATTAATTAATTAATTAATTAAT 2289
 Qy 1956 ATGGAAGTATACCAAGGAGGAAATGTTGAGGATCAAGATGATCAACGTAATTT 2015
 Db 2290 ATGGAAGTATACCAAGGAGGAAATGTTGAGGATCAAGATGATCAACGTAATTT 2249
 Qy 2016 AAAGCTGAAAACCTTCAATTAATTAAGGAGGAGTGTGTTCTGCAATTTTCTTCA 2075
 Db 2350 AAAGCTGAAAACCTTCAATTAATTAAGGAGGAGTGTGTTCTGCAATTTTCTTCA 2409
 Qy 2076 ATTAAGGAAATTTGACAGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 2135
 Db 2410 ATTAAGGAAATTTGACAGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 2469

Qy 2136 CAACAAATTAACATTTGACAGCGTTCAAGATGAGCAAGATTAACGACTTGTCAAAAAGTG 2195
 Db 2470 CAGCAAAATTAACATTTGACAGCGTTCAAGATGAGCAAGATTAACGACTTGTCAAAAAGTG 2529
 Qy 2196 GATTTAACCGTACAAAGATTAATTAATTTATTAACCAAAACCAATTAATGCTCTATT 2255
 Db 2530 GATTTAACCGTACAAAGATTAATTAATTTATTAACCAAAACCAATTAATGCTCTATT 2589
 Qy 2256 AATTTAATCTGTAATGCAAGGAGGAAATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 2315
 Db 2590 AATTTAATCTGTAATGCAAGGAGGAAATTTATTAATGTTTAAAGTTTAAAGTTTAAAGTT 2649
 Qy 2216 ACTTAAACAAATCAAGGCAATTTTACATTAAGCAACAAAGGCAACCAATTAAGCAATTT 2375
 Db 2650 ACTTAAATGATCAAGGCAATTTTACATTAAGCAACAAAGGCAACCAATTAAGCAATTT 2709
 Qy 2376 CAGCTTCCGACAAATTAACGAGTGTGATTAATGCAAACTTGAACGTAATGTCAT 2435
 Db 2710 AAACCTTCAATCAAGGCAATTAACGAGTGTGATTAATGCAAACTTGAACGTAATGTCAT 2769
 Qy 2436 TTAAGGATGATGCTCAATTTCTTAAACAAAGGCAATTTTGGCAACAAATTAAGGCA 2495
 Db 2770 TTAATGATGATGCTCAATTTCTTAAACAAAGGCAATTTTGGCAACAAATTAAGGCA 2829
 Qy 2496 GACAAAGGCAACAGTGAAGTGTGAAATGCACTTGAACAAATGCTTGAAGTACTTGA 2555
 Db 2830 GGGGAGGCAACAGTGAAGTGTGAAATGCACTTGAACAAATGCTTGAAGTACTTGA 2889
 Qy 2556 TTGCAAAATTAACGCTTAATTAACAGTGAAGTGTGAAATGCACTTGAAGTACTTGA 2615
 Db 2890 TTGCAAAATTAACGCTTAATTAACAGTGAAGTGTGAAATGCACTTGAAGTACTTGA 2949
 Qy 2616 TCAAAATTAAGGCA-----GTTGCGTCAATTAAGGCAACAAAGGCAACAGTGA 2669
 Db 2950 TCAAAATTAAGGCAACAGTGAAGTGTGAAATGCACTTGAAGTACTTGAAGTACTTGA 3009
 Qy 2670 GCGAAGCAATGCTTCAACAGTGAAGTGTGAAATGCACTTGAAGTACTTGAAGTACTTGA 2729
 Db 3010 GCGAAGCAATGCTTCAACAGTGAAGTGTGAAATGCACTTGAAGTACTTGAAGTACTTGA 3069
 Qy 2730 CAATTTACTTCAATTTTATTTGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 2789
 Db 3070 CAATTTACTTCAATTTTATTTGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 3129
 Qy 2790 GAGGCGATTAACATTAATTTCTGTCGCAACAGGCAACAAAGGCAACGCTTGAAGCA 2849
 Db 3130 GAGGCGATTAACATTAATTTCTGTCGCAACAGGCAACAAAGGCAACGCTTGAAGCA 3189
 Qy 2850 TTAATTTGTTGTTAAAGCAAGATTAATCAACGTTATCAAGTGAAGTGAATTTACTTTA 2909
 Db 3190 TTAATTTGTTGTTAAAGCAAGATTAATCAACGTTATCAAGTGAAGTGAATTTACTTTA 3249
 Qy 2910 GAAATGACACGTTGATGAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2969
 Db 3250 GAAATGACACGTTGATGAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3309
 Qy 2970 TTCCGCTGCTTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3029
 Db 3310 TTCCGCTGCTTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3369
 Qy 3030 CAAGGAGGCAACATTAAGGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3089
 Db 3370 CAAGGAGGCAACATTAAGGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3429
 Qy 3090 GAGGCAAGGAGGCTTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3149
 Db 3430 GAGGCAAGGAGGCTTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3480
 Qy 3150 CAAGGCTGTTAAAGGCTTGAAGGCAACCAATTAATTAATTAATTAATTAATTAATTAAT 3209
 Db 3481 CAAGGCTGTTAAAGGCTTGAAGGCAACCAATTAATTAATTAATTAATTAATTAATTAAT 3537

QY 600 TCTGAGCGGAGTTTGGCGAATGATCAAGCAAGGC---GACCAATTGGCGGTGCA 656
DB 956 AGTGGGATCAATGGTGGAAAGAGATATATATTAACCTTGGAGCTTACCTATGGA 1015
QY 657 TATCATATCTGACAGCTGCGAATACACAAATGAGCTGGAGCGATATGATTCG 716
DB 1016 GGTTCATGTTAATAGTGGAAATACCTTTGAAGATGGAACACAGCTGGTATACATTA 1375
QY 717 TATTTGGAGGCGATGTTGTTAAAGCGGAGAAATAGTTCATTAACGATTCAGGCTCA 776
DB 1076 GAATTAATGGCGAGTCAAAATTCCTAATTAATAGTCCACTACCTACGCGAGTTCA 1135
QY 777 AAGGGGACAGTGTCTTCCGATTTATTTATGATGCTGAAAACAAATAGTAAAT 836
DB 1136 TTGGGGATAGTGGTTCCTCAATGTTATTTATGATAGAAAGTTAAGAAATGTTATTA 1195
QY 837 AATGGATATTACGGGAAGGCAACCTTTTGAAGGCAAAATGGGTTCAATTGTT 896
DB 1196 AATGGCGTTTACGTGAAGGAATCCTTATGCTGACATGGAAGACGCTATCAATTA 1255
QY 897 CGCAATCTTATTTTGAATGAAATTTTGAAGATTTTCAATCATCACTTTACACCGA 956
DB 1256 CGAAAGATATTTTCAAGTATCTTAATCAAGACATTAAGCTAATTTTGGGATAT 1315
QY 957 GCTGTATGAGATGACAAATTAAGTGAATGATTTGTCAGGGGTC----- 1006
DB 1316 AATGCTAATATAGATTTATATATAGGAGTGCACAAATGAGAGTGGCAACATCA 1375
QY 1007 --TATATCTGAAATCAGATACATCAAGAAATTAATTAAGTTAGCAATATAGT 1064
DB 1376 AGTACATTAACCTTAAGAAAGCTATTCAGCTGAAGCAATAGGGTCTTTATGATTAATAGC 1435
QY 1065 TTACCTTTAAAGAGAGATTA-----GTTCAATCTAGATATGACGACCTAAT 1118
DB 1436 CAATCTCATGATGCTAGAGATTAATATAGCGATGATCTCCCTTATTAAGTCTAAT 1495
QY 1119 ATTTATTTCCAGTTTAAACAATGAGAAAGGCTAATTTTATGATCAAAAAGGA 1178
DB 1496 CCATGTGCGACATACATCATGAGAAAGATTTTACTTGGCAATGACAGAAAGGA 1555
QY 1179 TCATTAATCTTGACATCTGACATTAACAAAGGCGGAGTGTCTTATTTTGAAGTAT 1238
DB 1556 ACTTTAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1615
QY 1239 TTTACGATCTTCGAAT--TCTAACCAACTTGGCAAGAGTGGCATACATGTAACT 1295
DB 1616 TTTGTGTAAAGGCAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1675
QY 1296 GAAATAGCAACGTTTGAAGTAAATGCGGTGGAACATGATGACCTTCTAAATTT 1355
DB 1676 GAAAGAAATGATCTGTGAATGCGAGGTGCTAATTCAGAAAGCGATGCTTATCAAAATTT 1735
QY 1356 GGTAAAGAAATTTGACGTTCAAGCCAAAGGGAATTAAGGTTGATGATCGATAGGC 1415
DB 1736 GGGCTGGGAACCTTACTTGTATGTTAAAGGGAATTAAGGAAAGCTGAGTGTGCT 1795
QY 1416 GATGTAAAGTCAATTTGAGACAGAGGCAAGCATCAAGGCAACAAACAGCTTTAGT 1475
DB 1796 AACGTTTGTGTGTATGATCAACAAAGCAGATGATCAGTCAAAAACAAAGCTTTAA 1855
QY 1476 GAAATGGCTTGTGTTAGGCGAGAGGAGCTGTTCAATTAACATGATTAACATTTGAT 1535
DB 1856 GAAGTGGCAATGATAGGTGATGAGCTACGTTCAACTAAATAGTGAATCAAGTTGAT 1915
QY 1536 ACCGATTAATTTATTTGCGCTTTCGTGTGTGCTGCTTAAGTCTTAACGCGGATTCATTA 1595
DB 1916 CCTAAGAAATTTATTTGCGCTTTCGTGTGTGCTGCTTAAGTCTTAAGTGGGATTCATTA 1975
QY 1596 ACCTTTAACGTTATCCAAATTAAGGAGAGGAGGCAATGATTTGAACTAATTAATCAACT 1655
DB 1976 ACCTTTGAACGTTATCCAAATTAAGGAGGAGGATGATTTGGAACCAACAGCTTCT 2035

QY 1656 CAGCGCTAATGTCATTAATCTGAGACGAAGAGATTTGTTAATCTTAATGAAATTAAT 1715
DB 2036 CAACCGCAATTTATGATTTACAGGCAAGCAACTAATTAATCA-----GAT 2083
QY 1716 ATTAATTAATCTGATTAAGAAAGAAATTTGCTTAACAGGTTGGTTTGGCAACAGAT 1775
DB 2084 AGCAACCAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2143
QY 1776 AAAATTAACCAATGGGCAATTAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 1835
DB 2144 AAGCTTAACCAATGGTGTGTTAAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2203
QY 1836 TTGCTACTCTGAGTGTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1895
DB 2204 TTGTTGCTTTTGGGGGAGCAATTTAAACGCAATTAACGCAAAATAGTGTGATGTTA 2263
QY 1896 TTTTACGGGTAGACCGACACCGACGCTTCAATCATTTAATTAAGTTGTGTGAA 1955
DB 2264 GTTTTATGTTGCTTCAACGCTCATGCTTCAATCATTTAAGAAAGACCTTGTCTAAC 2323
QY 1956 ATGGAAGTATTAACAAAGGCGAATTTGTGGATCAAGATTTGATCAACGTAATTT 2015
DB 2324 ATGGAAGTATTAACAAAGGCGAATTTGTGGATCAAGATTTGATCAACGTAATTT 2383
QY 2016 AAGCTGAATTTCAAAATTAAGGCGAAGTGGGTGTTCTGCAATGTTCTTCA 2075
DB 2384 AAGCTGAATTTCAAAATTAAGGCGAAGTGGGTGTTCTGCAATGTTCTTCA 2443
QY 2076 ATTAGGAAATTTGACAGTCAAGCAATTAATGCAATTTGTTGTGTGTTGTTGTTGTTGTTGTT 2135
DB 2444 ATTAGGAAATTTGACAGTCAAGCAATTAATGCAATTTGTTGTGTGTTGTTGTTGTTGTTGTT 2503
QY 2136 CAACAAATTAACATTTGCAAGGCTTCAATGCAAGGATTAAGCACTGTCGCAAAAGTT 2195
DB 2504 CAGCAAAATTAACATTTGCAAGGCTTCAATGCAAGGATTAAGCACTGTCGCAAAAGTT 2563
QY 2196 GATTACGCAATTAAGTATTAATTTATTAATTTATTAATTAATTAATTAATTAATTAATTAAT 2255
DB 2564 GATTACGCAATTAAGTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTT 2623
QY 2256 AATTATCTGATTAATGCAAGGCAATTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTT 2315
DB 2624 AATTATCTGATTAATGCAAGGCAATTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTT 2683
QY 2216 ACTTAAACAATTAACAGGCAATTTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2375
DB 2684 ACTTAAACAATTAACAGGCAATTTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2743
QY 2376 CGACTTTCCGAAATTTCACTGCAAGGTTGATTAATGCAAGGTTAAGGTTAAGGTTAAGGTT 2435
DB 2744 AACTTTGAATTAACAGGCAATTTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2803
QY 2436 TTAACGATTAAGTCAATTTTCTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2495
DB 2804 TTAATGATCTGCTCAATTTTCTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2863
QY 2496 GAAAGGCAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2555
DB 2864 GAAAGGCAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2923
QY 2556 TTAACGATTAAGTCAATTTTCTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2615
DB 2924 TTAACGATTAAGTCAATTTTCTTAAAGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2983
QY 2616 TAAACATTAAGCA-----GTTGCGCTTCAATTAAGCAAGGCAATTTTAAAGCA 2669
DB 2984 TAAACATTAAGCA-----GTTGCGCTTCAATTAAGCAAGGCAATTTTAAAGCA 3043
QY 2670 GGAAGCAATGTTTCAACATTTGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 2729
DB 3044 GGAAGCAATGTTTCAACATTTGCAAGGCAATTTTAAAGCAAGGCAATTTTAAAGCA 3103
QY 2730 CAATTTACTCATCTTATTTGCTAATTAAGGCAATTTTAAATTAATTAATTAATTAATTAAT 2789

3104 CAATTTCTTCATCTTTATTTGGCTATTAAGGATTAATTAATATCAATGACCT 3163
2796 GAGGGCGATTACATTAATTAATCTGTCGCAACAGCGGAAAGAACCCGAAACCTTGAGAA 2849
3164 GAGGGCGATTACATTAATCTGTCGCAACAGCGGAAAGAACCCGAAACCTTGAGAA 3223
2850 TTAATCTTGGTGAAGGAAAGATATCAACCGTTATCAGATTAAGCTCAATTTACTTA 2909
3224 TTAATCTTGGTGAAGGAAAGATATCAACCGTTATCAGATTAAGCTCAATTTACTTA 3283
2910 GAAAATGACACGTTGATGACAGGTCATTAAGTTATTAATTAAGTAAGATGATGGCGAA 2969
3284 GAAATGACACGTTGATGACAGGTCATTAAGTTATTAATTAAGTAAGATGATGGCGAA 3343
2970 TTCCGCTTGCAATACCCATTAAGGACGAGAAATGCAATGATTTGTAAGACAGAG 3029
3344 TTCCGCTTGCAATACCCATTAAGGACGAGAAATGCAATGATTTGTAAGACAGAG 3403
3030 CAAGCAGACGACATTAAGGACGAGAAATGCAATGATTTGTAAGACAGAG 3089
3404 CAAGCAGACGACATTAAGGACGAGAAATGCAATGATTTGTAAGACAGAG 3463
3090 GAGCCAAAGTGGTCAAGAAAGAGCAGCGAGCGGTTCTGATACCTGCTGAT 3149
3464 AAGCCAAAGTGGTCAAGAAAGAGCAGCGAGCGGTTCTGATACCTGCTGAT 3514
3150 CAAGCCGTTAAAGCATTAAGGACGAGAAATGCAATGATTTGTAAGACAGAG 3209
3515 CAAGCCGTTAAAGCATTAAGGACGAGAAATGCAATGATTTGTAAGACAGAG 3571
3210 AGTAAGCGCAAAAACAAAAAGTGGCTCAAAAAGAGAGTGTTCCTGATCCCTGCTT 3269
3572 AGTAAGCGCAAAAACAAAAAGTGGCTCAAAAAGAGAGTGTTCCTGATCCCTGCTT 3631
3270 GATCAAGCCGTTGCGCATTAAGGACGAGAAATGCAATGATTTGTAAGACAGAG 3329
3632 GATCAAGCCGTTGCGCATTAAGGACGAGAAATGCAATGATTTGTAAGACAGAG 3682
3330 GAAAAGATGCTGATGCTCA---GAGAGCGGAGGAAACCAAGCAAAAGAAAGATG 3396
3683 GAAAAGATGCTGATGCTCA---GAGAGCGGAGGAAACCAAGCAAAAGAAAGATG 3742
3387 ATCAGCCGTTATCAATAGTGGCTTAATGAGATTAATGATGCAAGTAATGATGCTT 3446
3743 ATCAGCCGTTATCAATAGTGGCTTAATGAGATTAATGATGCAAGTAATGATGCTT 3802
3447 TCTGTTCAAGTAAATTAAGTGTCTTTGTAAGTCAAGCAATGCTGCTGAGCA 3506
3803 TCTGTTCAAGTAAATTAAGTGTCTTTGTAAGTCAAGCAATGCTGCTGAGCA 3862
3507 AATATCGAGAGATTAAGGACGAGTAATGATGAGTGTGCTGATGCAAGAGAG 3566
3863 AATATCGAGAGATTAAGGACGAGTAATGATGAGTGTGCTGATGCAAGAGAG 3919
3567 AAAAGCAATTAAGTGAATTTGGGTCGAAAGAAAGCTTAAGTGAAGCAATTTGGGCA 3626
3920 AAAAGCAATTAAGTGAATTTGGGTCGAAAGAAAGCTTAAGTGAAGCAATTTGGGCA 3979
3627 GTTTTCTGCAATGAGCTTCAATTAATCTTTGATGAACAGTTAAATATCAGCGACA 3686
3980 GTTTTCTGCAATGAGCTTCAATTAATCTTTGATGAACAGTTAAATATCAGCGACA 4039
3687 TTAAGCAATGAGCTTGGCTTTGCGCAATTAATCAATGAGTGAAGTGAAGTGAAG 3746
4040 TTAAGCAATGAGCTTGGCTTTGCGCAATTAATCAATGAGTGAAGTGAAGTGAAG 4099
3747 GTGGGAAAGGAAATCAGTGCAGTAATTAATGCTGAAGAAACCAAGCCGAAATTAATGCA 3806
4100 GTGGGAAAGGAAATCAGTGCAGTAATTAATGCTGAAGAAACCAAGCCGAAATTAATGCA 4159
3807 AAGGCAATTAATTAAGGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3866

4160 AAGGCAATTAATTAAGGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4219
3867 CAGCTTAATTTGAGAGTGAATGCTATTTATTAAGCGTGAATTAATCAATGAGGAA 3926
4220 CAGCTTAATTTGAGAGTGAATGCTATTTATTAAGCGTGAATTAATCAATGAGGAA 4279
3927 GTGAGAGTGAAGGAGCTTACGCTTGAATTAATGCTTAATGCTGAGATGAGTGAAT 3986
4280 GTGAGAGTGAAGGAGCTTACGCTTGAATTAATGCTTAATGCTGAGATGAGTGAAT 4339
3987 TATACATTTACTGCGAGATTAATCAAGCTTAAGCTTATTTCTGTAATTAATGTT 4046
4340 TATACATTTACTGCGAGATTAATCAAGCTTAAGCTTATTTCTGTAATTAATGTT 4399
4047 GATGTTCAAGGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4106
4400 GATGTTCAAGGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4459
4107 GAGCGTATTTGCAAAAAGAGTGGATTAAGGACGAGAAATTTACATTTCAAAATTC 4166
4460 GAGCGTATTTGCAAAAAGAGTGGATTAAGGACGAGAAATTTACATTTCAAAATTC 4519
4167 GCTTTATCTCAAAATCTCAAGTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4226
4520 GCTTTATCTCAAAATCTCAAGTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4579
4227 TTGGGCTATGCTGTTAAATCAAC 4252
4580 TTGGGCTATGCTGTTAAATCAAC 4605

RESULT 12
US-08-210-535-5
Sequence 5, Application US/08210535
Patent No. 5865424
GENERAL INFORMATION:
APPLICANT: Ambrosius, Dorothea
APPLICANT: Dony, Carola
APPLICANT: Rudolph, Rainer
TITLE OF INVENTION: Recombinant TGA Protease
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/210,535
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/820,701
FILING DATE: 10-Jan-1992
APPLICATION NUMBER: DE 41 00 704.2
FILING DATE: 11-Jan-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 41 40 699.0
FILING DATE: 10-Dec-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5965424man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: HUBR 1006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
 LENGTH: 4899 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 US-08-210-535-5

Query Match 3.8%; Score 164.8; DB 2; Length 4899;
 Best Local Similarity 58.0%; Pred. No. 3,6e-32;
 Matches 311; Conservative 0; Mismatches 222; Indels 3; Gaps 1;

QY 1152 CTATATTATGATCAAAAACAAGATCATTAATCTGCATCTGACATTAACCAAGG 1211
 DB 1202 CAATATGTTACCTTTGAGACACAGGCACTTGATTTGATTAATCAAAACATCAACAGG 1261
 QY 1212 GCGGGTGTCTTTATTTTGAAGGTAAATTACAGT--ATCTCAAAATCTAACCAACT 1268
 DB 1262 GCGGGGCGGCTGTTTTCAAGGGGATTCACAGTCAAAAGGCGCAATATGACATCACT 1321
 QY 1269 TGGCAGAGACTGGCATCATTAAGTAAATAGACACCTTACTTGGAAAGTAAAGG 1328
 DB 1322 TGGTTAGGTGCGGGATTTGATGCGGCAAAAAGTGTGGCGAAGTCAAAAT 1381
 QY 1329 GTGGAACATGATCGACTTTCTAAATTTGGTAAAGAACATGACGCTTCAAGCAAGG 1388
 DB 1382 CCGAATGGCGACGATTCGCAAAATTCGCAAAAGGCACTTTGAAATTAAGCGACAGG 1441
 QY 1389 GAAATTAAGTTGATCGATGAGGCGATGCTAAAGTCAATTTGAGACAGGCGAGC 1448
 DB 1442 GTTAAACCAAGGCAATTAAGTTCGCGACGCTACGCTTATCTGATCAAAAACCGAT 1501
 QY 1449 GATTAAGCGCAAAACAGCTTTATGAAATTTGGCTTGGTACGGCGAGGAGCTGTT 1508
 DB 1502 GCGCAAAAAGTTCAAGCTTCTCCCAAGTGGCATGTGCGGAGCGATGCAATTG 1561
 QY 1509 CAATTAACGATGATTAACATTTGATACCGATTAATTTATTTGGCTTTGGTGGT 1568
 DB 1562 GTATTAATTAATGTTCAATACATTAATCCCGATTAACCTAATTTGGTTTGGTGGG 1621
 QY 1569 CGCTTAATCTTACGGGCACTTCAATTAACCTTTAAAGTATCCAAATTCGACGAGG 1628
 DB 1622 CGTTGGATCCCATGCGCATGATGATCTTTGAAACATCCGCAACGTGAGTAAAGG 1681
 QY 1629 CCAATGATTTGTAACCAATTAACCACTCAAGCGCTATGCTATTTCTGCGGAA 1684
 DB 1682 GCGCGATTTGTCAACCAACACAGACAGCGCTTCCACATCACTTAACGAGTAA 1737

RESULT 13
 5268270-1
 Patent No. 5268270
 APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohner, Johannes
 TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
 NEGATIVE HOST CELLS
 NUMBER OF SEQUENCES: 6
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/171,872
 FILING DATE: 01-JUL-1987
 SEQ ID NO: 1:
 LENGTH: 4899
 5268270-1

Query Match 3.8%; Score 164.8; DB 6; Length 4899;
 Best Local Similarity 58.0%; Pred. No. 3,6e-32;
 Matches 311; Conservative 0; Mismatches 222; Indels 3; Gaps 1;

QY 1152 CTATATTATGATCAAAAACAAGATCATTAATCTGCATCTGACATTAACCAAGG 1211
 DB 1202 CAATATGTTACCTTTGAGACACAGGCACTTGATTTGATTAATCAAAACATCAACAGG 1261
 QY 1212 GCGGGTGTCTTTATTTTGAAGGTAAATTACAGT--ATCTCAAAATCTAACCAACT 1268

DB 1262 GCGGGGCGGCTGTTTTCAAGGCGATTAACAGTCAAAAGCGCAAAATTAATGATCACT 1321
 QY 1269 TGGCAAGAGGCTGGCATTAATGATGAAATTAAGACCGCTTACTGAAAGTAAATGCG 1328
 DB 1322 TGGTTAGGTGCGGGATTTGATGCGGCAAAAAGTGTGGCGAAGTCAAAAT 1381
 QY 1329 GTGGAACATGATCGACTTTCTAAATTTGGTAAAGAACATTCGATTCAGCAAGG 1388
 DB 1382 CCGAATGGCGACGATTTGCAAAATTCGCAAAAGGCACTTTGAAATTAAGCGACAGG 1441
 QY 1389 GAAATTAAGTTGATCGATGAGGCGATGCTAAAGTCAATTTGAGACAGGCGAGC 1448
 DB 1442 GTTAAACCAAGGCAATTAAGTTCGCGACGCTACGCTTATCTGATCAAAAACCGAT 1501
 QY 1449 GATTAAGCGCAAAACAGCTTTATGAAATTTGGCTTGGTACGGCGAGGAGCTGTT 1508
 DB 1502 GCGCAAAAAGTTCAAGCTTCTCCCAAGTGGCATTTGTGCGGAGCGATGCAATTG 1561
 QY 1509 CAATTAACGATGATTAACATTTGATACCGATTAATTTATTTGGCTTTGGTGGT 1568
 DB 1562 GTATTAATTAATGTTCAATACATTAATCCCGATTAACCTAATTTGGTTTGGTGGG 1621
 QY 1569 CGCTTAATCTTACGGGCACTTCAATTAACCTTTAAAGTATCCAAATTCGACGAGG 1628
 DB 1622 CGTTGGATCCCATGCGCATGATGATCTTTGAAACATCCGCAACGTGAGTAAAGG 1681
 QY 1629 CCAATGATTTGTAACCAATTAACCACTCAAGCGCTATGCTATTTCTGCGGAA 1684
 DB 1682 GCGCGATTTGTCAACCAACACAGACAGCGCTTCCACATCACTTAACGAGTAA 1737

RESULT 14
 US-08-232-463-14/C
 Sequence 14, Application US/08232463
 Patent No. 5670367
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 APPLICANT: SCHEIFLINGER, F.
 APPLICANT: FALKNER, F. G.
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 INVTU
 TELEPHONE: (703) 836-9300
 TELEFAX: (703) 683-4109
 TELEX: 899143
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7216 base pairs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2004, 17:29:18 ; Search time 28 Seconds
(without alignments)
4788.964 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273
Sequence: 1 MKKTVEFLNPLTACTISIGIV.....SQSGSLGKQONVGLGYRW 1394

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:.*
1: pirl:.*
2: pirl:.*
3: pirl:.*
4: pirl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	7273	100.0	1394	2	SG0762 IGA-specific serin
2	4178	57.4	1449	2	B81963 IGA-specific serin
3	4085	56.2	1457	2	DB1019 adhesion and penet
4	2634.5	36.2	709	2	CG4057 IGA-specific metal
5	1676	23.0	1431	2	A81018 serine-type peptid
6	1673.5	23.0	1541	2	A37023 IGA-specific metal
7	1662.5	22.9	1545	2	B41859 IGA-specific metal
8	1654.5	22.7	1561	2	SG1314 IGA-specific metal
9	1617	22.2	1532	2	A26039 IGA-specific metal
10	1609	22.1	1634	2	H54106 IGA-specific metal
11	1605	22.1	1702	2	A41859 IGA-specific metal
12	1555.5	21.4	1849	2	C41859 IGA-specific metal
13	1544.5	21.2	1815	2	C81169 IGA-specific metal
14	1536.5	21.1	1773	2	A81937 IGA-specific metal
15	1113.5	15.3	1377	2	IS4632 tsh protein - Bach
16	935	12.9	1366	2	SG5664 IGA-specific metal
17	847	11.6	508	2	SG1319 IGA-specific metal
18	845.5	11.6	497	2	SG1328 IGA-specific metal
19	836	11.5	508	2	SG1329 IGA-specific metal
20	835	11.5	508	2	SG1330 IGA-specific metal
21	824	11.3	486	2	SG1337 IGA-specific metal
22	824	11.3	486	2	SG1335 IGA-specific metal
23	823.5	11.3	507	2	SG1331 IGA-specific metal
24	821	11.3	508	2	SG1330 IGA-specific metal
25	819	11.3	496	2	SG1325 IGA-specific metal
26	811	11.2	496	2	SG1325 IGA-specific metal
27	806	11.1	482	2	SG1332 IGA-specific metal
28	707	9.7	462	2	SG1332 IGA-specific metal
29	678	9.3	471	2	SG1333 IGA-specific metal

30	604	8.3	1300	2	T00317 probable serine pr
31	368.5	5.3	384	2	B41500 IGA-specific metal
32	344.5	4.7	1569	2	A65044 hypotheical prote
33	341.5	4.7	1528	2	D85912 hypotheical prote
34	341.5	4.7	1571	2	C91068 hypotheical prote
35	326.5	4.5	1430	2	AF0357 probable autotrans
36	313	4.3	949	2	T08658 hypotheical prote
37	292.5	4.0	3705	2	AD0123 probable autotrans
38	288	4.0	836	2	D64988 yeJ0 protein - Bac
39	283.5	3.9	1536	2	A43855 high-molecular-we
40	281	3.9	836	2	A98014 hypotheical prote
41	281	3.9	836	2	C85858 hypotheical prote
42	280	3.8	1477	2	B43855 high-molecular-we
43	268.5	3.7	1327	2	B90674 Aida-I adhesin-lik
44	268.5	3.7	1349	2	E85524 probable beta-bar
45	258.5	3.6	1645	2	JN0896 crystalline surfac

ALIGNMENTS

RESULT 1

SG0762
IGA-specific serine endopeptidase (EC 3.4.21.72) precursor - Haemophilus influenzae (str N)/Alternate names: adhesion and penetration protein hap
C/Species: Haemophilus influenzae
C/Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 08-Dec-2000
C/Accession: S60762; S78301
R/St Gene III, J.W.; de la Morena, M.L.; Falkow, S.
Mo. Microbiol. 14, 217-233, 1994
A/Title: A Haemophilus influenzae IGA protease-like protein promotes intimate interacti
A/Reference number: S60762; MUID:95131744; PMID:7830568
A/Accession: S60762
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-1394 <STC>
A/Cross-references: EMBL:U11024
A/Experimental source: strain N187
A/Accession: S78001
A/Molecule type: protein
A/Residues: 27-33 <ST2>
C/Genetics:
A/Genes: hap
C/Superfamily: IGA-specific metalloendopeptidase
C/Keywords: hydrolase; serine proteinase
F/-26/Domain: signal sequence #status predicted <SIG>
F/27-1394/Product: IGA-specific serine endopeptidase #status experimental <MAT>
F/243/Active site: Ser #status predicted

Query Match 100.0%; Score 7273; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKKTVEFLNPLTACTISIGIVSQSMAGHYFGIDVYRPAENKKEFYVGNQKYYNNQ	60
DB	1	MKKTVEFLNPLTACTISIGIVSQSMAGHYFGIDVYRPAENKKEFYVGNQKYYNNQ	60
QY	61	GOLVGSMTKAPMDIFSVSRNCAVALVENGYIVSAHNVGYTDVDFGAEENNDDHRT	120
DB	61	GOLVGSMTKAPMDIFSVSRNCAVALVENGYIVSAHNVGYTDVDFGAEENNDDHRT	120
QY	121	YKIVKNNYKXNDLHPEDYENRLEKFTVTEAPDIMSNNNGSTYDRKYPERVRIG	180
DB	121	YKIVKNNYKXNDLHPEDYENRLEKFTVTEAPDIMSNNNGSTYDRKYPERVRIG	180
QY	181	SGRQFRNDQKQOVAGAYHTLGNTHNRGAGNGSYLGDDYRKAGEYGPPIINGSK	240
DB	181	SGRQFRNDQKQOVAGAYHTLGNTHNRGAGNGSYLGDDYRKAGEYGPPIINGSK	240
QY	241	GDGSGPMFYDAEKQKWLINGIIRREGNPFBEKENGFPOLVRKSPFDEIFERDLHTSLYTRA	300
DB	241	GDGSGPMFYDAEKQKWLINGIIRREGNPFBEKENGFPOLVRKSPFDEIFERDLHTSLYTRA	300

QY 301 GNGVYTTISGNDGQSSITQKSGIDPSEIKITLAMSILPKKXDKVHNPRYDGNITVSPILAN 360
 DB 301 GNGVYTTISGNDGQSSITQKSGIDPSEIKITLAMSILPKKXDKVHNPRYDGNITVSPILAN 360
 QY 361 NGRTIYFMDOQKQSSILPASDINQAGGLYEENFTVSPNSQOTWQAGIHVSENSTVTKW 420
 DB 361 NGRTIYFMDOQKQSSILPASDINQAGGLYEENFTVSPNSQOTWQAGIHVSENSTVTKW 420
 QY 421 VNGVEHRLSKIKGKGLTHVQAKGKNGKSSISVGGKVTILEQADDOQKQAFSEILVSGR 480
 DB 421 VNGVEHRLSKIKGKGLTHVQAKGKNGKSSISVGGKVTILEQADDOQKQAFSEILVSGR 480
 QY 481 GTVQLNDKQPDYDKFEYFGSGRLDNGHSLFKRIQNTDEGAMIVNHNITQAAVYTT 540
 DB 481 GTVQLNDKQPDYDKFEYFGSGRLDNGHSLFKRIQNTDEGAMIVNHNITQAAVYTT 540
 QY 541 GNESTIVPNGNNINIKLDYRKELIANGFSGTDKNKNGRLMLYKPTTEDRTLLSGGTN 600
 DB 541 GNESTIVPNGNNINIKLDYRKELIANGFSGTDKNKNGRLMLYKPTTEDRTLLSGGTN 600
 QY 601 LKGDITOTKGLFEFGSPTPHAYNHLNKRSEMEGIPOGELIWMOHMINTKAEINPOIX 660
 DB 601 LKGDITOTKGLFEFGSPTPHAYNHLNKRSEMEGIPOGELIWMOHMINTKAEINPOIX 660
 QY 661 GGSAAVSRNVSIEGKMTVSNNAATPGVVPQONTICTRSDWTGLTTCQKVDLDTKVI 720
 DB 661 GGSAAVSRNVSIEGKMTVSNNAATPGVVPQONTICTRSDWTGLTTCQKVDLDTKVI 720
 QY 721 NSIRKQINSINLTNATAVKGLAKNNTLTINHSQPTLSNNTQIGTRISDNGSTA 780
 DB 721 NSIRKQINSINLTNATAVKGLAKNNTLTINHSQPTLSNNTQIGTRISDNGSTA 780
 QY 781 TVDNANLNGNVLITSAQPSLKNSHFOIGKGTVTLLNATWTPSPTTLQMLTLN 840
 DB 781 TVDNANLNGNVLITSAQPSLKNSHFOIGKGTVTLLNATWTPSPTTLQMLTLN 840
 QY 841 SITLNSAYASASNNTPRRRSILETETTPSAHRFNLTIVNGKLSGGQFQGTSLSPYK 900
 DB 841 SITLNSAYASASNNTPRRRSILETETTPSAHRFNLTIVNGKLSGGQFQGTSLSPYK 900
 QY 901 SDKIKLSDNDEGYLISVNTGKEPTELEDTLVESKNOPLSKIKFLTBNDHVDKAL 960
 DB 901 SDKIKLSDNDEGYLISVNTGKEPTELEDTLVESKNOPLSKIKFLTBNDHVDKAL 960
 QY 961 RYKLVKNDGEFRILNFIKEQELHDIYRAQAEFTLEAKQVEPILACTQYGEKVSRRPA 1020
 DB 961 RYKLVKNDGEFRILNFIKEQELHDIYRAQAEFTLEAKQVEPILACTQYGEKVSRRPA 1020
 QY 1021 RAAPFTLPPQSLNALBEAKQALTAETQSKAKTKVRSKRAVPSDPLLDGSLPALBA 1080
 DB 1021 RAAPFTLPPQSLNALBEAKQALTAETQSKAKTKVRSKRAVPSDPLLDGSLPALBA 1080
 QY 1081 LEVIDAPQSEKORLAQEBEAKOKKODLISYSNSALSLSATWSMISVODELDFV 1140
 DB 1081 LEVIDAPQSEKORLAQEBEAKOKKODLISYSNSALSLSATWSMISVODELDFV 1140
 QY 1141 DOAASAVMTNIAQDKRRYSDAFRAYQOQKTLROIQYOKALNKHGIVFSHSRDNTF 1200
 DB 1141 DOAASAVMTNIAQDKRRYSDAFRAYQOQKTLROIQYOKALNKHGIVFSHSRDNTF 1200
 QY 1201 DBQVKNHATLTWMSGFAYQWGDIOFGVNVGIGISASMAEBSRKIRKALNIVGNASY 1260
 DB 1201 DBQVKNHATLTWMSGFAYQWGDIOFGVNVGIGISASMAEBSRKIRKALNIVGNASY 1260
 QY 1261 QPRLGOLGQPYGVNRYFIRENYGSEERKVTSLAEKNVNAKIRVUTTEPTDINIV 1320
 DB 1261 QPRLGOLGQPYGVNRYFIRENYGSEERKVTSLAEKNVNAKIRVUTTEPTDINIV 1320
 QY 1321 KPEFFVNVVSNNAVQTTNLTVLOQPFGRYQKXGKALILHFOISAFISKSQSGOL 1380
 DB 1321 KPEFFVNVVSNNAVQTTNLTVLOQPFGRYQKXGKALILHFOISAFISKSQSGOL 1380
 QY 1381 GKQONVGVNLGYRW 1394
 DB 1381 GKQONVGVNLGYRW 1394

DB 1381 GKQONVGVNLGYRW 1394
 RESULT 2
 B81963
 Iga-specific serine endopeptidase (BC 3.4.21.72) NMA0457 [imported] - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #ext_change 02-Feb-2001
 C:Accession: B81963
 R:Parthill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 J.; Holroyd, S.; Jørgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: AB1775; M01D:2022356; EMBL:10761919
 A:Accession: B81963
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1449 <PAR>
 A:Cross-references: GB:AL162753; GB:AL157959; NID:97379120; PIDN:CA883754.1; PID:9737920
 A:Experimental source: serogroup A, strain Z2491
 C:gene: iga2; NMA0457
 C:Superfamily: Iga-specific metallopeptidase
 C:Keywords: hydrolase; serine proteinase
 Query Match 57.4%; Score 4178; DB 2; Length 1449;
 Best Local Similarity 56.2%; Pred. No. 4.6e-196;
 Matches 816; Conservative 214; Mismatches 325; Indels 98; Gaps 17;
 QY 10 FLTACISLGIWSCWAGHTYFGIDYQYRDPANRKGKFTVGAQIKYVNGQGLVTSMT 69
 DB 27 YLALCLSGELLQAMAGHTYFGINQYRRDPAENKKEVAGKALEYVNNKKGLVSKMT 86
 QY 70 KAPMIDSVSRNRYVALYENQYIVSAHNVGYTDVDPGAEQNDQRFYKIVRNNY 129
 DB 87 KAPMIDSVSRNRYVALYENQYIVSAHNVGYTDVDPGAEQNDQRFYKIVRNNY 146
 QY 130 KKNLHPEDDYNPRLHKVTEAPIDMTSMNGSTSDRTKPERYRISGRCRMD 189
 DB 147 KPNSHPTNGYHNPRLHKVTEAPIDMTSMNGSTSDRTKPERYRISGRCRMD 206
 QY 190 QDKXQVAGAYHTTAGNTHNORGANGYSTLGGDYAKAGEYPLPIKSGKSGSPMT 249
 DB 207 DDHEDLISYSGAWLIGGNTHQMGANGVSLSDVRYHANDYGPMPIGAAGSGSPMT 266
 QY 250 YDAKKQKMLNGILRENPBEKENGFOYVAKSYF-DLPEEDLITSLIYAKNGYTTIS 308
 DB 267 YDKNNKMLNGVLOTGYPSGREGFQLRKQWFFYDDIYRGDHTYFEPKRSNGHRSPT 326
 QY 309 GNDGGSITQSGITPSEIKITLAMS-PLKEXDXYHNPY--DQPNITSPRLNCE 363
 DB 327 SNNNGYITETNKVSNPKLQYVRLPDESLEITDK--EPVYAGGVNQVQPRLNCE 384
 QY 364 TLVFMDOQSSILPASDINQAGGLYEENFTVSPNSQOTWQAGIHVSENSTVTKW 423
 DB 385 NLSPFDYNGKLLISNNINGAGGLYEPGDPVSPNNETWQAGVHISEDSTVTKW 444
 QY 424 VEHDRLSIKIGKGLTHVQAKGKNGKSSISVGGKVTILEQADDOQKQAFSEILVSGR 483
 DB 445 VANDRLSKIGKGLTHVQAKGKNGKSSISVGGKVTILEQADDOQKQAFSEILVSGR 504
 QY 484 QLNDDKQPDYDKFEYFGSGRLDNGHSLFKRIQNTDEGAMIVNHNITQAAVYTT 543
 DB 505 QLNDDKQPDYDKFEYFGSGRLDNGHSLFKRIQNTDEGAMIVNHNITQAAVYTT 564
 QY 544 SILPENGNNIKLDYRKELIANGFSGTDKNKNGRLMLYKPTTEDRTLLSGGTN 603
 DB 565 SILPENGNNIKLDYRKELIANGFSGTDKNKNGRLMLYKPTTEDRTLLSGGTN 624
 QY 604 DLTOTKGLFEFGSPTPHAYNHLNKRSEMEGIPOGELIWMOHMINTKAEINPOIX 663
 DB 625 NITQNGKLFPSGTPHAYNHLNKRSEMEGIPOGELIWMOHMINTKAEINPOIX 684

RESULT 4
 C64057
 IGA-specific metalloprotease type 1 homolog - Haemophilus influenzae (strain Rd K72)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 31-Oct-1997
 C:Accession: C64057
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
 R:Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.
 R:D.M.; Brandon, R.C.; Fife, L.D.; Fritchman, J.L.; Fuhrman, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.T., McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Ventter,
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: C64057
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-109 <TTGA>
 A:Cross-references: GB:U34710; GB:I42023; NID:91573200; PID:91573211; TTGR:H0248

```

Db 121 IQYKRNKNYQAMKREKPRPDGDYHMPRLKHTVEAEVGTITKMGKYTYAREVPERVNI 180
QY 180 GSGCQFMRNDQDKGD-----VAGAYHYLTAGNTINRGAMGYSYLGADVKACEYCP 233
Db 181 GSGGQYRTDKDBETWVSHSYYSVCAVRYLLTAGNTHTGCMGNGTVMLSGNVSPHYCP 240
QY 234 LPIAGSGDGSMSMFLYDAKQKMLNGILREBNPEEGKENGFOYVAKSYF-----284
Db 241 LPTGSGDSDGSMSMFLYDAKQKQMLNAAVQJTHPEFGKMGFOILREHFWNEVLAVDI 300
QY 235 DEIFERDLHSLYTRAGNGVYLTSGNMGSGSIT-----QKSGIPSEKITLAMSILP-LK 339
Db 3C1 PSVFOR-----YTPINGHTSVSNNDGSGKULLTRBPKDSSPRAKSEVGTWKLFPNSLN 354
QY 340 EKDYKHPNRYDGNITSPEPLANKEELTFMDQKQSLIFASDINOGAGLYPEGNFTVSEN 399
Db 355 QTAKEHKAAGAIVNIQPFMEYKRIYLGQDQNGTTLTIENNINOGAGLYPEGNFTVYXK 414
QY 400 SNO-TWOGAGIHVSENSYTYKRWGVHEHRLSKITGKTLHQAQKENGKSLSYGDSKYL 458
Db 415 QNNLTWQAGAVSIGQDQATVEKVNHPENDRLSKITGTLTVNGKGMKLSLMAQNKYLL 474
QY 459 EGOADPGNGQASEGLVSGKGTVOJLNDKQKPDTPKFPFGGSGRLDLNGHSLTEFKIC 518
Db 475 DQADHBAGQKQAKXEGVIGSRATVQJLNSDQYDPRNIYFGRGGRDLNGHSLTEFKRI 534
QY 519 NTDEGAMIVAHNTTQAAVNTITGNEBSIVLP-NGNNINKLDYAKEJAVNGMFGSETDKKN 577
Db 535 NTDEGAMIVAHNTTQAAVNTITGNEBSITLASNKNKINLKJLSKEJAVNGMFGSETDKKN 594
QY 578 GRMLUYKPTTEDRITLLSGGTNLKGDITOTKQKLPFSGRPTPEAHVNLKMRSEKGI 637
Db 595 GRMLUYKPTTEDRITLLSGGTNLKGDITOTKQKLPFSGRPTPEAHVNLKMRSEKGI 654
QY 638 QGEIYMHDMINTEKAEHFOIKGSAVVRNNSITEGMVSNNAKATGCVEN 692
Db 655 QGEIYMHDMINTEKAEHFOIKGSAVVRNNSITEGMVSNNAKATGCVEN 709

```

[illegible]


```

Db 303 LFLSSYDPAAGNKKSSMOEWNITKSOPTKDVANKSASLIGSKTIDYSSMSKGTSTTIG 362
Qy 310 NDNGQSGITOKSGIPSEIKITLANNMLPLKEXOKVHNERYDGPNTYSPPLANGETLYPMD 369
Db 363 GB-----KSLNVDLAD-----GKDK-----NKGKSYTF-- 386
Qy 370 QKQSSLLFASDINGAGGLYFEGNFYSPNS-NCWQOAGIHSSENSYTTWKNVAGYEHDR 428
Db 387 EGSGETLNNNNIDQAGGLFPEGDYEVKGSNTTWKAGVSAEKKTYTWKVNHPQYDR 446
Qy 429 LSKIGKGLTVYAKGKNGSISVGDGKYLIEQOADDQGNKQAFSEIGLVSAGTYOJLND 488
Db 447 LAKIGKGLTVYEGTJUNGSLSKVDGTVILKQGTNGSG-QAAPSAGVYSGSTLYVAND 505
Qy 489 KQPTDXYEFGFEGSLDINGSLTFPRKIQNTDEGAMTVNNTTQAAVNTTIGNESIYLP 548
Db 506 KQVPSNITFYGFRGSLDINGSLTFPDHRIINIDGAPLVNHNMTNNSNITITGESLITDP 565
Qy 549 N----- 549
Db 566 KTTTPYIDAPDENPYAFRRIKDGGQLYINENTYVALRKASSTRSELPRNSGESNEN 625
Qy 550 -----GNNINKLDYRKEIAYNGWFGSTDKNHNGRNLNLYKPTTEDRTLLSG 597
Db 626 KLYMGKTSDEAKNWNHNNHNERMNGFNQYFGE-BEGKNNQNLNVTFKKSQKQNFLLTG 684
Qy 598 GTNLKGDITQTKGLTFPSGRPLPPIAVNHLKNGWSENGI-----POGEIYWDHWD 647
Db 685 GTNNGDLTVYEGKGLTFPSGRPLPPIA-----RDIAGISSTKDPKPEANNEVVEVDWD 726
Qy 648 INRTEKAKNFOIKGSAVVS-SNVSSIEGNTWYNNNAKATGVVPNQOITICRSDTGL 706
Db 737 INRNFKATMNTVGNASLSGSENVANTTSNTLSNKAQVHIGY-KTGCTVYVRSDYIGY 794
Qy 707 TTCQVLDLDTKYNSIPETQINGSLNLTNNATANKVLAJNANVTLTNH8OPTLSNNA 766
Db 795 YCTCTDKSD-KALNSFNTLRGNVALTSESANF-VLKANLTFETIQSGNSQ----- 845
Qy 767 TQIGKIRLSDNSTATYDNNANNGVNHITDSAQSLKNSHPSHQIQDGKGTYYTLNATWT 826
Db 846 -----VRLTENS-----HMHJLTGN-SDVHQJDLNANH----- 871
Qy 827 MPOSTTLQNTLNNSTTLNLSAYSSANNPRBSLETETTPTSABERFETLTVNGKLSG 886
Db 872 -----IHNSA--DNSNNT-----KNITLTVN-SLSG 896
Qy 887 QCTPQTSLSFGYKSDKLKLSNDAEZYIISVANTGKEPRTLBOGLTVYSSKNOPLSDKL 946
Db 897 NGSFYIITLDSNKGQKVVTKSATGFTLQVADKTCGEPMH-NELTLFASAKAQ--RDHL 953
Qy 947 KFTLENDHVDAGLRKLVKANDGFRFLANIKYQBELANDV----- 987
Db 954 NVSLVGNVTVDIGAMKXKJANVNGRYDLNP-EVERNOQVDTNNTITPNK-QADVPSPV 1011
Qy 988 -----RAEQERLTLEAKOVEPT----- 1004
Db 1012 SNNEIARVDEAVPPAPAPSPSETTETVAENKQESKTVEKHOQOAITETIAQNRVAK 1071
Qy 1005 -----AKTGTG-----PRVSRRAA 1620
Db 1072 AKSNVYKANTQTNNEVAGSGSETKETQTTETKETAIVEKEKAKVETKQSEVKRVSQSP 1131
Qy 1021 RAAPFDTLPOS-----LLNMLEAK-QABELTQKSAKAKYKRSKRAVAPDPLD 1071
Db 1132 KQOGESEVYQOAPAPRENDPTVNIKEPOSQTTTADTQTOPAKETS-----SWEQPYTE 1185
Qy 1072 QSLFALBALEVIDAFOQ-----SEKDLAQBAAKQKOK----- 1107
Db 1186 ST--TVWNGSVVENENTTATQTPVNSSESNKPKNRHRSVRSVPHVVEPACTSSND 1243
Qy 1108 -----DLISRYNSALSLSA-----TVNSLSTVQDELRLTV-DQOASAVNTNIAQD 1154

```

```

Db 1244 RSTVALCDLDTSTNNINAVLSDAPAKQAFVALNKGAVSQHIISQLENNRQGVNVAWSNTSM 1303
Qy 1155 KRIYSDAPFAYQOQKTNLRQIGVQKALANG-RIGAVFHSNNTDEQVKNHATLTM 1213
Db 1304 MKVYSSQYRRFSKSTQ--TOLGMDQTTISNNVOLGAVFTYVNSNNDKATSKRN-TLQOV 1361
Qy 1214 SGPAQY-----QWEDLOGVVNGTGISAKKAEQSKIRHKAINVGNVNASYQRLQGLGI 1269
Db 1362 NRYSKTYADNHW--YIGIDLGVGFQSGKIQTNHNAKFAKHQAQGLTAGKAPENLGNPQI 1418
Qy 1270 QPYFGVNRPIERENTYQSEBEVKTPSIAFNRYAGIRVDYTFPTDNISVKEPFVNTV 1329
Db 1419 TPIYGVRRYSYNSNADFALDQARITKYNP:SVYTAFAQVDSLTYT-HLGBFSVTEILSARY- 1476
Qy 1330 DVSANAVQTYNLTVLVQOPFGKRYQKVEGLKAEILFQJSAISQSGSQLGQGVNCGK 1389
Db 1477 DAAQSGSKNTVNGVDPAYVNERQOYNAGLTKYHNWKLISGLITAKQAEKQXTAEIK 1536
Qy 1390 LGY 1392
Db 1537 LSF 1539

RESULT 7
B41859
IGA-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - Haemophilus influenza
C/Species: Haemophilus influenzae
A/Variety: Strain HK393
C/Date: 04-Mar-1993 #sequence, revision 18-Nov-1994 #ext_change 08-Dec-2000
C/Accession: B41859
J/Poulsen, K.; Reinholdt, J.; Kilian, M.
J. Bacteriol. 174, 2913-2921, 1992
A/Title: A comparative genetic study of serologically distinct Haemophilus influenzae ty
A/Reference number: M41859; MUID:92234949; PMID:1373717
A/Accession: B41859
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-1545 <POU>
A/Cross-references: GB:M67490; NID:G148908; PIDN:AAA24967.1; PID:G148909
A/Experimental source: strain HK393
A/Note: sequence extracted from NCBI backbone (NCBI:P_97283)
C/Suprafamily: IGA-specific metalloendopeptidase
C/Keywords: hydrolase; metalloproteinase

Query Match 22.9%; Score 1662.5; DB 2; Length 1545;
Best Local Similarity 28.6%; Pred. No. 3.6e-73;
Matches 482; Conservative 246; Mismatches 511; Indels 447; Gaps 59;

Qy 6 FRNLPLTACISLS:VSGAMGHTTGGIDYQYRDPANRKGFTVGAONIKYVNRQQLVG 65
Db 6 FKLNFTALVAVALLPYTEALVVDVDYQIFRDFANRKGKTSVGATVVEVRDKNNRPIG 65
Qy 66 TSWTKA-PRHDPESYS-RNGVALVYENQYISVNH-NVGTIVDVG-AEGN---APDOH 117
Db 66 NVLPNGIPMIPDSVVDVKRIATLVNPPYVGVGVHVSNGVSELSFENLNGMNNKNAKHA 125
Qy 118 RFT-----YKIVKNNY-----KCONLHPYEDDYHNPRLKEFTMAPID-MTS 160
Db 126 RDISSEENRYIVKNEYEPYKNGKAVTTEDQAQKREDOYMPRLDKVTVVAALTEASTD 185
Qy 161 MNGSTYSNRYKYPERVYRIGSGRF-----KNDQ-DKQOVAGAHYTLAHTN 209
Db 186 SSTAGTYNNKQKQYFYFRLSGTGYENGTRYELMWGKEGQKSD--AGGVNKLGVGNAY 243
Qy 210 NQSGAGNGYSY-----LGGDVAKAGEY-GPLPIAGSK-----GSGSPMEIYDAK 254
Db 244 TYGIAGTPYVNHENHNDGLIFGNSNNEYINPKELSKKPLTYAVLVGSSGSPFLVYDRK 303
Qy 255 QKMLINGILBEGNPFEGKE-NGFQVRSYPTDE:FEEDLMTLSI-----YPAQNG-VYT 306
Db 304 GKMLPLGSYDVWAGVNNKQWQWNTIKYKPEFAKTIYEQYSGSLIGSKTDYMSNGSTST 363
Qy 307 ISGNDNGQSGITOKSGIPSEIKITLANNMLPLKEXOKVHNERYDGPNTYSPPLANGETLY 366

```

```

Db      364 ITGGE-----KSLANDLAD-----GKDP-----NHGSVT 389
Qy      367 FMOQKGSLLPASDINOAGAGLYFEGNFTVSPNS--NQTWQAGIHVSENSTVWTKNGVE 425
Db      390 F--ESSGILLNNNDQAGGLFFBSGDIYVNGTSDNTYMGAGVYAGKATVTKMKNQ 447
Qy      426 HDRLSKIGKGLTAVQAKGENKGSISVGGKVLLEQADQDNKQAPSEHIGVSGRGVQL 485
Db      448 YDLRLATIGKGLTIVEGTGDKGSLKAGDSTYLLKQNTSSG--QHPASVGVISGRSLV 506
Qy      486 NDDKQPTDDEYFEGRGRLDNLGSLTPKQIQTDBGAMVYNNHTTAAAVYTTGNEST 545
Db      507 NDDKQVDPNSIYFGRGRLDNLGSLTFEDIRNIDGALVYHSTSGKSTVYITGDNLI 566
Qy      546 VLPNG----- 550
Db      567 TDBNNVSIYKPLJEDNPALQITQYQULYNEENKTYTAKADASIRSEFPNGRS 626
Qy      551 -----NNIKLDYRKELAYNGMFGETDNKNGRLNLYKPTTERTL 594
Db      627 NNSMLYMGTEKADQKAMNHINNEBNGFNGYFGE--BEKNNGNLNTVFKGSEQRRL 685
Qy      595 LSGSTLKGDITQTKGLFSGRPPTHAYNLAKNGSEAGI-----PQGEIWD 644
Db      686 LTGGLMNGDLNVQGLTFLSGRPTHA-----RDIGLSITKDSHPSENNEVYE 737
Qy      645 HDNINPFAENFQKGSAAVVS--RVVSSIEGNVTNNNMAFVGVEVNOQNTCTCSPM 703
Db      738 DDMINRFKATINVTNNATLYSGRVVSTSTNTAANNKHHIIG--KAGDTVCVRSY 795
Qy      704 TGLITQKVDLTDEKVINSPKQINGSLNLTDAATVYKGLAKINGVLTJNHSQPTLS 763
Db      796 TGYVTCPTDKSD--KALNSNPNTNLGNVNLTESANF--VLGKANLFGTIGRNGS-- 849
Qy      764 NNATQGNRLSDNSTATVDMANLGNVHLTSAQFSLKNSHSHQIGDGKTVTLNNA 823
Db      850 -----VALTENS--HMHLTGN--SDVQLDLANGH----- 875
Qy      824 TWMPDPTTLQKLTNNSTLINSAYASANNTPRRSLTETPTTSAEHRFTLVYNGK 883
Db      876 -----IHLNSA--DNSNNVT-----KYNLTLYN-S 897
Qy      884 LSGGSEFQPTSSLFQYSDKUKLSDNABDYLLSVRYTKEPELTTEQLTVESKONGPLS 943
Db      898 LSGGSEFYITDLKSKQDQKVVVTKSATGNFTLQVADTGEPRH--NELTLFDASKAQ--R 954
Qy      944 DKLKFTLBNDDVDAALRYKLVKADDEPRLNPIKEGLANDLV----- 987
Db      955 DMLNVLGNVTDLGAMKYKLRVNGRYDLNP--EVEKRGQVDTVTNITPNNICADVP 1012
Qy      988 -----RABQERTLEAKOVEPT----- 1004
Db      1013 SVPSNNEIARDEKPPAPAPATPSETTETVAENSKQSEKTYERKQADLETTRQREY 1072
Qy      1005 -----AKTQGE-----PKYRSR 1017
Db      1073 AKEAKSVKANTCTNEVAGSGSEKETEOTTEKETATVEKEKAKVETEKEQEPKYSQ 1132
Qy      1018 RAARAAAPPTLDQ-----LNLALAK--QALTLPTQSKAKTKKVSKAAYSD 1068
Db      1133 VSPQKQSETPQAPARENDPTVNIKEPQCTNTTADVEQPAKETS-----SVEQEP 1186
Qy      1069 ILDDSLFALEALFETAPQO-----SEKRLAQEAEKQKQK----- 1107
Db      1187 VTST--TYWTGNSVNEPENTPATTPQTVNSESNNKPRHRBSVRSVPHNVPATIS 1244
Qy      1108 -----DLISYNSALSSELIS--TVNSMLSVODELRLFV--DQAGAVTNI 1151
Db      1245 SNDSSTVALCDLTSTNTVAVSDAKAQFVALNKGKAVSOHLSOLEMNEBQYVWVSN 1304
Qy      1152 AODRRYDSAFRAVQOKNTLRQITGVAKALNG--RIGAVSHSBSDNTEFDQVYNNATL 1210

```

```

Db      1305 TSNMNTSSQYRRPSSSTQ--TQLGMDQITSNNTQLGCVTTVYRRSNNFDKATSKA--TL 1362
Qy      1211 TWMSGFAQY-----QWGDLOFQVNVGTISASKKAEROSKIRHRKAINVYNASVQFRLQ 1266
Db      1363 AQVNVYSKYADNMH---YLGDIGYCKPQSKLQNNNAKARHTAQFGLTGKAPNLGN 1419
Qy      1267 LGIQPYFQVNRVYFERENYQSEEVAVTSPSLAFNRNNGIRVDYTFPTDMISVYFV 1326
Db      1420 FGITPIYGVARSYLSADPFDQARIKVNPISVKTAFQAVDLSTY--HLGEFSVTPILISA 1478
Qy      1327 NVVDYSNANCTVMTLVLOQPFGRWQKEVGLKZELHPIQSAFISKSGSGLKQGNV 1386
Db      1479 RT-DKMGSSKINVNGVDYFNVENQQYNAIGLKIKYHNVKSLIGGLKAPAEKOKTA 1537
Qy      1387 GWKLG 1392
Db      1538 EKLKSP 1543

RESULT 8
S61314
IGA-specific metalloendopeptidase (EC 3.4.24.13) homolog - Neisseria meningitidis
C:Species: Neisseria meningitidis
A:Variety: Hfl3
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 08-Dec-2000
C:Accession: S61314
R:Lombolt, H.; Poulsen, K.; Mogens, K.
Mol. Microbiol. 15, 495-506, 1995
A:Title: Comparative characterization of the iga gene encoding Iga1 protease in Neisseria
A:Reference number: S61314; MID:95502561; PMID:7783620
A:Accession: S61314
A:Status: preliminary
A:Residue type: DNA
A:Residues: 1-1561 <LOM>
A:Cross-references: EMBL:X82474; NID:q732873; PDB:QAA57851.; PID:q732874
C:Superfamily: IGA-specific metalloendopeptidase
C:Keywords: hydrolase; metalloprotease

Query Match      22.7%; Score 1654.5; DB 2; Length 1561;
Best Local Similarity 28.8%; Pred. No. 8.9e-73;
Matches 494; Conservative 239; Mismatches 508; Indels 473; Gaps 55;

Qy      1 MKKTVFLNPLTACISL--GIVSAMAGHYFGIDYQYRPAEKKKKFTVGAQINIKYV 58
Db      1 MKTKFKNALSLSTIFLAYALPYSEALVVDVQIFRPAENKGFVVGATDLVKN 60
Qy      59 KQGLVGTSTKAPMLDFSVGRN--GVNALVENQYIVSAHV-----GYTV 105
Db      61 KQGNIGNALSNDPMDIDFSALVYKRIATVVDQYAVSYKAKAEVHTFYGGYGNHNV 120
Qy      106 DFGAGNNPDGKEFTYKIVKRNKY-----KNIHFYEDDYENRLLAKVTEADIDMT 159
Db      121 -----ADKEN-----EKRYVEQNNYEPHKAWSASNLRLR--DYNNARFRNKEVTEVADTPT 170
Qy      160 SNNG--STYSRTYRPRVRISSQ-----FWNDDKQ--DQAAAGHYTLAG-- 206
Db      171 DAGGIDYTKKNNFSSFVRVAGKQLVYEKGAHVQENKQYDLIDSQAYVYALAGTP 230
Qy      207 -----NTHNQDAGN--GYSYLGGVYKAGEVGPLPIAGSKGSGSPMFITYAEKQ 255
Db      231 YKDINDQNTTBGLIFSNHHTHSABELKQALSODALTNVGVLTGSSGPLAFKPKKN 290
Qy      256 KWLINGILRENGPEKEN--GPOLVRKSYDEIFERDLHTSLYTRANGVYTTGNDNG 313
Db      291 QWFLGTYDVWAGY--GKSGWQMNITKKPEADKIKORD-----NA3-----TLXGVGH 338
Qy      314 QSGITQKSGIPISEIKITLNMSPLEKDKQHNDRDGNISPRLNKSETLVPMQKQK 373
Db      339 HMKCTGTSIHGSTAVLAN--NEED-----ANKQKVTF--ENN 375
Qy      374 SLIFASDINQAGAGLYEENFTV--SPNSQWQAGIHVSENSTVWTKVNGVHDLRSKI 432
Db      376 TLVLDQNTINQAGSLPFKQDITVYKANNNDITVLGAGIDVADQKTVWQVKNPNDRLAKI 435

```


Dd		1241	ADAVSTNTMSALSDMASQTOSTL-----LDTGAYLTRHIAQSRPADAEKRVYWMENSTGYG	1295
Oy		1156	KRVSDFANVQQOCTNLRIOGVQALA-NGRIGAFVSASNDTPEQRYKHATLTMMS	1214
Dd		1296	RDAVASAQYRRFSRSRTQ-TOIGIDRSLSENMOIGGLTYSOSHPTD-QAGKRTFVOAN	1353
Oy		1215	GFAQYQWGLQCF-GVNNGVGISASVMABEQSKHKHAINGVNASVQGPFAGLGIGIOPYF	1273
Dd		1354	LKGKYLVLMNAVVDAGTGAGSGRSLPCTQQAANNFTSIOTELTLGNITKLINOPEIVPSA	1413
Oy		1274	GWNRYTEFEENVQSEEVRYKTPSLAFNNYNAGRIVDYTTPTDNISVKPY----EPVVNY-	1328
Dd		1414	GIRKSLSADYKIGDDSVYSSMAVKTLLNGLDPAVR-KKNGLTYKPLLSAAFYANV	1472
Oy		1329	---VDVSNANVQTVNLVTVLQDPFGRYQWEKWLKAPILHPDISAFYSKS--OGSOLGKO	1383
Dd		1473	KGVNVVGKGSFPXYAADN-----QQQSYAGVALLYRWLVNLNGTSIKGKOLEKO	1521
Oy		1384	QNVGVNLGYRM	1394
Dd		1522	KSQGIKIQLTF	1532
 RESULT 10 H64106 IGA-specific metalloendopeptidase [EC 3.4.24.13] type 1 - Haemophilus influenzae (strain N).Alternate names: immunoglobulin A1 proteinase type 1 C.Species: Haemophilus influenzae C.Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #ext_change 08-Dec-2000 Accession: H64106; A11500 R.Fleisschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kjerlakege, R.; Gockayne, J.D.; Scott, J.; Shirley, R.; Liu, L.-I.; Glodet, A.; Kelley, J.M.; Weidman, J.M.; Brandon, R.C.; Fine, L.D.; Freilichman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995 A.Authors: Ghem, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, F.O.; Venter, A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A.Reference number: A64000; MUID:95350630; PMID:7542800 A.Status: nucleic acid sequence not shown; translation not shown A.Molecule type: DNA A.Residues: 1-1694 <TTR> A.Cross-references: GB:U032779; GI:I42023; MID:g1574009; PIDN:AAC22651.1; ZID:g1574019; A.Experimental source: strain Rd KW20 R.Gundy, F.J.; Plant, A.G.; Wright, A. Infect. Immun. 58, 320-331, 1990 A.Title: Localization of the cleavage site specificity determinant of Haemophilus infl A.Reference number: A41500; MUID:90129281; PMID:2105270 A.Accession: A41500 A.Molecule type: DNA A.Residues: 1-377 <GRU> A.Cross-references: GB:X59800 A.Experimental source: strain Rd KW20 A.Note: the authors translated the codon TGG for residue 319 as Thr C.Function: A.Description: this proteinase is classified as type 1 because it cleaves at a proline C.Superfamily: IGA-specific metalloendopeptidase C.Keywords: hydrolase; metalloproteinase Query Match 22.1%; Score 1609; DB 2; Length 1694; Best Local Similarity 26.5%; Prod. No. 1.7e-70; Mismatches 536; Indels 542; Gaps 52; Matches 479; Conservative 251;				
Oy		6	FRLNFLACISIGIVSQAMAGHTYEIDYQYRDPAENKGFKEVGAONIKUYNKQGOJVg	65
Dd		6	FKNFILMTVAALPTTBALVRDVDYQGFDPFNKGRSRVCAITVEFRDKNHSLG	65
Oy		66	TSMTKA PMIDPSVVS-RNGVVALENQYTYSVAH-NVGTVDVDC-AEGN----	NPDGE 117
Dd		66	NVLFNIGIPMIDPSYVDYDKRIATLNPQYVGVGVNSNGSELHGMLNGNMNNGAKSH	125
Oy		118	RFT-----YKIVKENNV-----KKCNLPEDYDHNPRLHKFVEAAPID-MTS	160

```

Db      126 RVSSSENRFPSEKNEIPTKNGKATVTDQQRREDIYMPRLDKFTVTEPALEASTA 185
Qy      161 NMNGSIYSDRTKYPERVIRIGSGRQFWNRDQDKDQ-----VAGAYHYL 203
Db      186 SSDAGTYNDQNKTPAVRLGSGSQP---IYKGDVNSLLINHHVGGNNLKLVGDAYYQ 242
Qy      204 TPAKTHQKAGNGGYILG-----GVRAAGYGPLPLASSKDDSGSPMETIYDEKGM 257
Db      243 IAGTPYKVNENNGGLGFENSKREHSDPKGILLQDPLTNVAVLGDSSPLFYDEKGM 302
Qy      258 LINGLIEBGNPFEGKE-NGPOLYKSYFDEIFPBDLHTSLYTRAGNVTTISGNDGGS 316
Db      303 LFLSYDPFAGTKKSGWEMNITYKEPAKTVLDKDTAGSL-----TGS 345
Qy      317 ITQSGGIPSEIKITLANNMPLKKEKDVHNPYD-GPNISYPRLNGEITLYMDQKQSL 375
Db      346 NTQVMMPTGKTSTVSNQS-----ESLNDVLPDSQDPTDKKNNHCKSVTL--RGSGL 397
Qy      376 TPASDINQAGLXFEENFTYSPNSNO-TWQAGLHSEBNTVMKNGVEHDLSTK 434
Db      398 TLNNNTQAGAGLFFEGDYEKGTSDSTTKKAGSVADGKVTWKVHNPCKSDRLAK 457
Qy      435 GTLHYQKGRNGKSGISVGDGVILEQDADQCKQAFSEIGLVSGRTYQVLDNDQFPTD 494
Db      458 GTLHYEGKNGKSLKVGQGVYLLKQADANNKVAFGQVIGVSGRTYVLDNDQVDPN 517
Qy      495 KYEFGPRGLDNGSLTPYKIONTEGAMVNTTQANNTTTSNBYLEN----- 549
Db      518 SIYFPGRLDANGNNLTFEIRINIDQARLVNHTSKTSTVITSESLTDPPTITPY 577
Qy      550 ----- 549
Db      578 NIDAEEDNPAFRIRKQGGQLYNLNNTYVALRKASTSELPKNGSENNMLYNGK 637
Qy      550 -----GNINKLKEIAYNGFGETDKKNGKGLMLIYKPTEDRTLLSGSTNGK 633
Db      638 TSDAKXNMWMLHNERMNGFYGE-BEGKNGMLNVTYKGSSENRILTGSTNG 696
Qy      604 DTTQTKGLPSPGRTPPAHYN-HLNKSEMEKGIPOGLYMDHMTNTPPAENFOIGK 661
Db      697 DKTVEKGTLEPISGRPTPHARDIAGISSTKQDHPAENRNVVEDWLNRFATKINTVN 756
Qy      662 GSAYVS-RVNSIEGWTVSNANAATGVVNPQONTICRSDDMTGLTQCKVDLETKVI 720
Db      757 NITLISGRVANITSTIASDQAKVHIGY--KAGDVCYRSDYGYTTCTTDKSD-KAL 813
Qy      721 NSIPTQINGSIMLTNATANYKGLAKLNGAVTLTNHSGFTLSNNATQIGNRLSDNSTA 780
Db      814 NFPNATVSGNVVLSONANF-VLQKAMLFGISGTSNQ----- 851
Qy      781 TYDNANLNGVHLTDAQSLKNSHFHQIGKGTVTYLENATYMPSTTLQVLTLLAN 840
Db      852 -----VRLT-----BNSEHMLTGSQSVNQLINDK 875
Qy      841 STITLNSAYSASNNTPRRBSLETPTPTSAEREPNLTIVNGKLSGGGTQFSSLPFYX 900
Db      876 GHILNANDAKVTT-----YNTLTV-SLSGSGSYTILTDSKQ 916
Qy      901 SDKLTSLNDAGGYILSVNTEKPEPTEQLTLVESKQNPQSLDKFLTLNDHVDAGAL 960
Db      917 GKXVTVKSAKNGFTLQVADKTGER-TKNLTLFDA-SNATNNMLVSLVGNVTYDLGAW 973
Qy      961 RYKLVNDGERLHNLKQELANDV----- 987
Db      974 KYKLVNNGRYDLYN-EVEKXQVDTYNTITTPNNTQADVPSPNNEELARVETVP 1011
Qy      988 -----RAQAPRTLE-----ARQVPTAK--TQTCG-- 1011
Db      1032 PAPAETPSETTEVAENSKQESKVEKQEDATETTAQNGVAEAEAPSKANTQTNVEA 1091
Qy      1012 -----PKYRSRRARAAP-----DTL 1028

```

```

Db      1092 QSGSETEETGTTZEIKETAKVEKEKAKVEKDEIQEAPQMASETSPKQAPKVESTDK 1151
Qy      1029 PQOSLINAL-----EAQABLTAFQOKSA----- 1053
Db      1152 VEETQVQAPQTOSTTVAAAEATSPRSKAPRETQSPKTNAPVTVVSKXQENTTDQ 1211
Qy      1054 -----KTKVRSKR-----AVPSDPLEDQ----- 1072
Db      1212 TEREKTAKEITEKQEPPOVASQSPKQSETYQPOAVLSENPVTYNNAEVYQALQ 1271
Qy      1073 -----SLFALAEVVIDAPQ----- 1089
Db      1272 QTSATVSTQAPAPENSINTGSAATLETAKSKDPQETASTEDASGHKANTVADNSVA 1331
Qy      1090 -----SEKRLAQEP-----AEQCKQK----- 1107
Db      1332 NSESSDPKRRRRSISQPETSAEETTAETASTEDTTLADSKRSPNRSRVSREPTV 1391
Qy      1108 -----DISRYSNALSLSA-----TVNSMLSVODELRLFY-DOAKSAWT 1149
Db      1392 TNSDSYVALALDLTNTNTNAVSDMAKAQFVALNKGAVSGHISQLEBNNBEGQVYVW 1451
Qy      1150 NIAQDRRYDSDAPRAYQOKTNLROIGVQALANG-RIGAVFSHSRSDNTFEDQVKNHA 1208
Db      1452 SNTSMENENSSQVRRFSKSTQ-TQLGMDQTTISNNVQLGQVFTVRRNSNPD-KASISN 1509
Qy      1209 TLTMSGFAQY-----QWGDQFGVNVGTGISASQMAEBSRKTHKALNVGNVASYQFL 1264
Db      1510 TLAQVNFYSKYADNHW---YLGIDIGYKFGQSMKTNHAKFAHHTQFSLTGKAFNL 1566
Qy      1265 GOLGIQPYGVNRYFIRENYQSEBEVAKVETPSLAFNRYNAG-RVDTPTPTPDN-SVAPYF 1324
Db      1567 GNFGITPTVGRYISYLSNANFALADRIKKNPISYKTAAPQVLDSTY-HIGESVTPIL 1625
Qy      1325 FVNYVDVSNANVQTVNLTVLQOPFGRYQKEVGLKAEILHFOISATISQSGSLGKQ 1384
Db      1626 SARY-DTNQSGSKLNVQYDFAVYVENQOYVNAGLKIKHYNVLTLIGLTKAQAEKOK 1684
Qy      1385 NNGVKG 1392
Db      1605 YAEIKLSP 1692

```

```

RESULT 11
A41859
IGa-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - Haemophilus influenzae
C:Species: Haemophilus influenzae
A:Variety: strain HK715
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Dec-2000
C:Accession: A41859
R:Poulsen, K.; Reinholdt, J.; Kilian, M.
J. Bacteriol. 174, 2913-2921, 1992
A:Title: A comparative genetic study of serologically distinct Haemophilus influenzae ty.
A:Reference number: A41859; MUID:92234949; PMID:1373717
A:Accession: A41859
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1702 <POU>
A:Cross-References: GB:M87489, NID:G148306, PIDD:AAA2966.1; PID:G148307
A:Experimental source: strain HK715
A>Note: sequence extracted from NCBI backbone (NCBI:P_97282)
C:Superfamily: IGA-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

```

```

Query Match      22.1%, Score 1605, DB 2; Length 1702;
Best Local Similarity 26.4%, Prid. No. 2.7e-70;
Matches 479; Conservative 231; Mismatches 536; Indels 550; Gaps 52;

```

```

Qy      6 FRLNFTACISIGVSAWAGHTYFQIDQYRDPFAENKGFYVAGNRIKYTNKQGLVQ 65
Db      6 FRLNFTALVVAALPTTEBALVYRDVDYQIFRDPFAENKGRSFGVATVWEVFDKNNHSLG 65
Qy      66 TSMTKA-PMIDFSVVS-RNSVNAALVENQIVSVAH-NVGTVDDPQ-ABGN---NEDQH 117

```

Dh	66	NVLPGSITPILDFSVVDVDRXIAITLINDQYVGVSVNGVSEJLHFNALNGMNNKGDKXH	125
Qy	118	RFL-----YKIVERNY-----KKONLHFYEDDYNHPLKFTVTEAPID-ATS	160
Dh	126	RUVSSEBRNYSVAKOBYPTKLNGKAVTEEDQORRRBDYVMPRLKFTVTEAPIDASTA	185
Qy	161	NNNGSYSDRKTVEYRVRJGSGRQPMRNDQKDO-----VAGYHFL	203
Dh	186	SSPAGYIVDQKRPAPAFRRJGSGQF-----YKKGDVYSLINNHVEGNNLKVGDVATYTG	242
Qy	204	TACNTHNQCGKNGSVYLG-----GPRKAGEYPLFIAGSGDGSQPMVYDAKOKM	257
Dh	243	IAGTPYKNNHNNKLJGFNGSKSEHSDPKILSDBPLNNYALJGDSGSLPYVDRKKGX	302
Qy	258	LINGILREGNPECKE-NFQIVKRSYFDEIFERDLATSLYTPDAGNYTTISGNDGQGS	316
Dh	303	LFGSGYDFWAGYNGKKGXWQGMNLYKPEFAKTVLDKDTAGSL-----TSS	345
Qy	317	ITQKSGIPEIKILITANMGLPIKEDDYVHPRD-GPIYSPRLNGETLFFMOQKQSL	375
Dh	346	NTQYNNPLFGKTSYVSGS-----ESLNDLFDSSQDTPSKKONHKGSVTL--KSGSL	397
Qy	376	IFASDINQAGGLYFEGNFTVPSVSNQ--TWQAGIHSVSNVYTWKGVNHEDELSTKX	434
Dh	398	TANNIDQACGLFFBEDYEVKTSJDSITWKAGVSVADGKVTWYKHNPKSDRLAKIK	457
Qy	435	GTLHQAGKENGKSVSGCKVILLEGQADQGNKQAFSEJGVSGSGVTOJLNDKQDPD	494
Dh	458	GTLIVGKENGKGLKVGKGVILKQDQDANKKQVAFQVGVISGTVLVNDKQVDN	517
Qy	495	KPYFGRGGRJDLNKGSHLFXKILQNTDEGAMVYNNHTYQAAVTTIGNESIVLEN----	549
Dh	518	SYFGRGRRLDANGNLTFFEHIRLIDGARLVNHNSTKSVTLTIGSELITDENTITPY	577
Qy	550	-----	549
Dh	578	NIDAPDEDPYAFARLIDGQOLYLNLBNYTYALRKGASTRSELPRNSGENSENMLYMK	637
Qy	550	-----GNINKLYRKEIAYVGMFGETDQKNGKGNLKYKFTEDRLLSGTNLKG	603
Dh	638	TSEAGAVNMMHINNEKNGRFGYFG-EGKGNKLNVTFKKSKQNRFLITGNTLNG	696
Qy	604	DITQYKGLFSGSPPEPHAV--HINKRSMEGIPQGEIYMBDHMINFTPAENFOJNG	661
Dh	697	DLAKEGTLFLSGRPTPHARDIAGISS-KXQDHFANNEVVEDDMINKPFAKINIVYN	756
Qy	662	GSANVS-RUVSSIGMWTSSNNANNTGCVPNQNTICTRSDMGLTQCYVLDLTKYI	720
Dh	757	NATLYSGRNVANITSNLTASDAKVVHGY--KAGDVYCVSDYGVYTTTQDKSD-KAI	813
Qy	721	NSLPIKQINGSLINTDANATYKGLAKINGVLTNNHSCFTLSNNAUTOIGIRLSDNSTA	780
Dh	814	NSFANVNSGNNLISGANP-VLGRKALFGTISGTONQ-----	851
Qy	781	TVEKAMINAVLTDASQFSLKSHSFHQIQGKGYTTLLENKATYMBEDTTLQNTLNN	840
Dh	852	-----VALT-----ENSHMELTGDNSVQJNLDX	875
Qy	841	STTLNSAVSASNNUTRRRSLETTPTSAHERPNTLVNGKLSQGFQPTSSLGXY	900
Dh	876	GHIHLNRQNDANKYT-----NYTLTN-SLSNGSFYIYLDLSNKG	916
Qy	901	SDKLKLSNDBEGYILSVNTGKPEPTEJQTLTVESKONQPLSLKFTLBNHYDAKAL	960
Dh	917	GDKVYVTKATSNFTLQVAKTGEF-TKHELTLEFDA--SNATRNUNLAVSLVGNFTVDSAM	973
Qy	961	RYKLYKNODEPFLNPIKBEOLANDY-----	987
Dh	974	KYKIRAVNERYDLYNP--EYKRNQTVDTNITTPNIIQADVPYSVNNHEIAVETPPV	1033
Qy	988	-----BAGQERTLE-----AKOVEPTAK--TGTGE--	1011

[illegible]

Db 6 FKNFALITVAALPTTEALVBDVDYQIFRDPAEKKGKSVGATNEVRDKKNSQ 65
Qy 66 TSMTKA-PMIDSVVS-RNGVALVENCITVVAH-ANGTDDVDFG-ABGN---NDQ 117
66 SALPNGLPMIDSVVDVKRLATILVNPQYVGVKAVNSGSEILFGLNKNMNNNGAKSH 125
Qy 118 RFT-----YKIVKNNYKKNHHPY-----EDDYNNPLAKETLAPID-MT 159
Db 126 RDUSSERNRYTVEKNNPTENVTSTTKEBODAQKREDTYMLDKETVAPIASI 185
Qy 160 SNNKSTYSDRTKPEPRVITSGKOF-----RNDQDYGQ 195
Db 186 ANNKGETNSDKYPAFVRLGSSGQFLYKKSRYQLITTEKQGLNMDVGGDLLEL 245
Qy 196 VAGAYHLLTAGNTHNOKAGNGSYLG-----GVRAAGEYPLFIAGSKGSSGMEI 249
Db 246 VGNAYTGLAGTPYKNNHNNGLIFGNSKEHSDPKGLISQPLNYAVLADSSGSLFV 305
Qy 250 YDAEKQKMLINGLREBNPFEGKENGFLYKRSYFDE--IFEDLHLSLYTAAKNGYTT 307
Db 306 YDREKGMV-----FLGSDYFVAGNKKSNQEMNLYKHEPAKTYQYKSGSLTG 355
Qy 308 SGND-----NGGSTITQSGIPESEIKITLANSPLKEKQVNPYDOPNYSRLANG 362
Db 356 SNTQYTWQATGSTITTGGEPLSVDLT-----DKD-----KNNHG 392
Qy 363 ETLFPMQKQSLIFASDINOAGAGLFEENFTVSPNSQ--TWQAGIHVSNTWTYKY 421
Db 393 KSITL--KSGGITLNNHIDGAGGLFPEBDEIVKGTSDSTTWAGAGVVAAGKVTWYK 450
Qy 422 NGVEHDLSTKIGKGLHYQAKGKSGISVGGKVLLEQADQKQKQAFSEIGNSRG 481
Db 451 HNPXYDLAKIGKGLVVBGKGNBGLKVGCGVILKQKADANNVQAFSGYGVYSRS 510
Qy 482 TVQANDKQPDITKTFEFGGRGLDNGHSLTFKQNTBEGAMTNHTTQAANVTITG 541
Db 511 TLVANDKQVDPNSITFEGRGRLDNGSLTFDHRKIIDGARYNNHNTVTSITITG 570
Qy 542 NESIVLN-----549
Db 571 BSLLTNPTVITTSYINBAODDHPRLRSIPRYQVFNODNRSYTILKAGSTRSLPONS 630
Qy 550 -----GNINKLDYRKELIANGFGETD--KXKNGKRLNLYPTTED 590
Db 632 GSNSEMLYMGKTSDAKRYVMHINERANGNGFGEERTKLTQKGLNVTENGSKDQ 690
Qy 591 KRLISGGLTKGDITQTKGLFSSGPRPHAYNHLNKNRSGMEGI-----PGE 640
Db 691 KRLITGGLTNGDLAVEKGLFLSGRPHY-----RDLIGISSTKXDPHTTENE 742
Qy 641 TWQHDATNRTYAEHFQKGSAAVVS-RNVSSIBGNVTSSNANAIFGVWPKQNTICT 699
Db 743 VVEDDMINRFAITMTNTGMAISYSGRVNATISNTITASMAQVHIG--KTDPYCV 800
Qy 700 RSDWGLTTCQKVDLDTFVINS:PKTQINGSLNLDNATNMGAKLNGVTLTNHSQ 759
Db 801 RSDYGYTGENSNLSB-KALNSNFPNTNLGANNLTENS-----833
760 FTLNVAOTGNTB:SNSTATNTNANNGVH:TLNDAOFSLKNSHSHOIGSDKSTVT 813

Qy 980 -----QELNDYVR-----AEQAEHTLE-----397
Db 1020 APSAQSNNEBIARVETPPPPAPATASIASEQETPRPAETLQAPAMEETNTANSITIAK 1079
Qy 998 -----AKOVEET--1004
Db 1080 SPTATQENPNSSVSEETTERKVAENPPQENETVAANEQEAETEPQNGVAKQOPTVE 1139
Qy 1005 AKTOYGE-----1011
Db 1140 ANTOQNETAQSEKTEBTETAKSEPTESVTSENQPEKTVSOSTEDKVVKEKEKAV 1199
Qy 1012 -----PKYSRBARAA--FPDILPDQSLINALEAKQAEETQKSKATKVRSK 1061
Db 1200 ETEETQKAPQVTSKEPPKQAEPAPEEVPDT--NAEBAQALQOPTVAAETTSPNK 1257
Qy 1062 RA-----VFSD-----ELLQD-1072
Db 1258 PAEBTQPESEKTNAPETPVVSENATQPTETETAKVEKENTQEVQVNASQSEFKQEP 1317
Qy 1073 -----1072
Db 1318 AAKPQATQKPAEPARENTLTNANGEPQAPQOPQUSAVPTTGETAANSKAPAPQAO 1377
Qy 1073 -----SLPAA 1079
Db 1378 AKQETPARENVS-TWTKEPQOTSATVTHQPAKETSSNVQPAPENSINTGSATTME 1437
Qy 1080 ALEVIDAPQO--SEKOR-----1094
Db 1438 TAEKSDKQMETVENDROPEANTVADNSVANNSESSKSRRSVSQPEKISAETTIV 1497
Qy 1095 -----LAQEBE-----KQKQ 1106
Db 1498 ASQOETTVDSVSTPKERSRTRRSGVQNSYEPVELPTNNAEMENVSANNVANSQPAL 1557
Qy 1107 KDLISRYNSKALSELB-----TVNSKLSVQDELDRLFV--DQAGSAVNTIADQKRRYS 1160
Db 1558 RNLTSKNTNAVLSNAVAKAQFVALNIGAVASQHLQLEBANNEGQYNWVISSTNKNKYS 1617
Qy 1161 DAFRAYQOQKTNLQOIGVQKALANG-RIGAVFSHSSDNTPEQVKNKAAITLMMSGFAQY 1219
Db 1618 EQYRREPSKSTQ-TOLGMDQTTSMNVQIGVFTYKRSNNPD-KASSNTTLAQVNTYSKY 1675
Qy 1220 -----QMGDLQGVNNGVGLISASMAEBOSEKIHKALNVCNASSQVFLGQGLQYFQV 1275
Db 1676 YADNEM--YIGIDLGYGQSYLQTNNAKFAAHTAQJLKGAFPLGFAFKPVTGV 1732
Qy 1276 NRYFIRENYSSEKRYKTPSLAFNRNYAGIRVDTTPTDNI SVKPYEFVNYVDVSNAN 1335
Db 1733 RYISYSNADPALADRLKNPISVKTAPQVDLSYTY-HLGESTITPLSARY-DAMQGN 1790
Qy 1336 VQTTNMLTVLQOPFGRYWQEVGLKAEILHFQISAF-SKSQSQGLQKQOONGVYGLAY 1392
Db 1791 GKINVSVDYFAYINENQOQNAGLKLYHNKVLISLIGLITXKQAEKQKTEVGLST 1847

QY 1027 -----TLDPQSLINALEKQ-----AELT----- 1045
 Db 1193 EKKAEELAKORAEARENOALAAARAKAEAEKQAEALAEARQAEAKAEALSAQAKA 1252
 QY 1046 -----AETQ----- 1049
 Db 1253 AERQALAAQOKALAEQAEAREKAAELAVKQAEETERTAEALAKORAAABAARQAEARQ 1312
 QY 1050 -----KSKAKTKKV-----RSKRAVS 1066
 Db 1313 TAELEARQAEERQAEALSAQAEETDRAEASAKRAEERERQAAOSQORKKRAAPQ 1372
 QY 1067 D-----PLDPQSLFALAEAL----- 1081
 Db 1373 DYMAASQNRPRKRGKRETLAPRPSQFSSAYAPRALINPWPENDEYHPLDALEDEN 1432
 QY 1082 --EVIDAPQSEKD-----RLAQEAE-----KQKQKDLISRSNSA 1117
 Db 1433 VSESVDTSDKQPDNTEHEKEVETSLQPRAPQPRQAAATOPQAAQAAQDAVSTYNTSA 1492
 QY 1118 LSELSTYNSMLSVODELD-----RLPVDQASAWNTNLAQDKRYDSDAPRAY 1166
 Db 1493 LSPAAASTQSTL-----LDTGASLSTRHQAQSRADAEKQSVMSNIGYGRDYASQYKAF 1547
 QY 1167 QOQKTLRQIGVOKALA-NGRIGAVFSHSRSDNTFDEOVYNHATLTPWAGPQYOWADLQ 1225
 Db 1548 SSKRTQ-TQIGIDRSLSENNQIGVYYSQSHFD-QASGKNTFVQANLKGKYLNDAM 1605
 QY 1226 F-GVNVGTGISASRAEBSRKIRKAINGVANSTQFRLQGLQIPYGVNRYFLEREN 1284
 Db 1606 YMAGDIGASLRSLQYQKAMFNRTSIQTLGTLKMOFEELVPAGIRYSRSLSSAD 1665
 QY 1285 YQSEHYRYKVTGSLAENNYNAGIRVDYTFPNDNIVKPY-----FVNVY-----VDV-SNNA 1335
 Db 1666 YKQNDVSVKSVMSKTLTGDLDPAYR-KVGNLTPRLLSAIYANIGKGVNAGNSF 1724
 QY 1336 VQTTVNLTVLQPPRGYQKEVGLKAEILHFOISAFISKS-QQSOLQKQONWELGYR 1393
 Db 1725 VYRADN-----QQQYASGALLYRNTLVNNGSITKQKLEKQSGQIKQIR 1772
 QY 1394 W 1394
 Db 1773 F 1773

RESULT 15
 154632
 tsh protein - Escherichia coli
 C:Species: Escherichia coli
 C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 08-Dec-2000
 C:Accession: 154632
 R:Provenance, D.L.; Curtiss, R.
 Infect. Immun. 62, 1369-1380, 1994
 A>Title: Isolation and characterization of a gene involved in hemagglutination by an avian
 A:Reference number: 154632; MUID:94178945; PMID:8132344
 A:Accession: 154632
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1377 <RES>
 A:Cross-references: GB:U27423; NID:9469235; PTDN:AAA24698.1; PID:9469236
 C:Superfamily: Iga-specific metalloendopeptidase

Query Match 15.3%; Score 1113.5; DB 2; Length 1377;
 Best Local Similarity 26.2%; Pred. No. 1.9e-46;
 Matches 403; Conservative 197; Mismatches 519; Indels 421; Gaps 59;

QY 3 KTVRLNF-LTACISLIGIVSQAMAGTTFGIDYQYRDPAAENKGFVGAONKIVYKOG 61
 Db 29 KSVRLCFPVLLILIPVLSAGSLAGTVNELGQLFRDPFAENKGMFRPATITAIYKOG 88
 QY 62 QLVGTSMTKAPMIDSVV-SKNGYALVENQYTVVAVNNGYTDVDFGAEKGNPQDRFT 120
 Db 89 EFVGT-IDKAMMPDESADVSEIGVATLINDQYIASVKINGGTYNVSFG-DGEN-----R 140

QY 121 YKIVKNNYKONLHPYEDYENRHLKFTAEADMTS-NNNGSYDRTKYPERVRI 179
 Db 141 YNIDRNNAF-----SLDFHAPRLDKLVEVAPTAVTAQAVAGAY-JDEKRVFVYRL 193
 QY 180 GSGRFWRNDQKQVQVAGATHILTGNTNORAGNGYSLGGDYRKAG-----EY----- 231
 Db 194 GSGTYIKDSNCSQLQWKGAYSWLTGGT-----VGSLSYQNGEMISTSSGLVFDYKLN 247
 QY 222 GPLFIAGSGDSSPMFIYDAEKQKMLINGILRENGPEKENGFOLVKSYFDEIF--E 289
 Db 248 GMPFIYEGADSSSLPFAFDIVQKKNVYGVLTANGAGRGKNNKAVIILDIQSKPFED 307
 QY 290 RDLHTSLYTAGNGYTTISGNDNQSSITQKSGIPSEIKITLAWMSFLKEDKVHNPY 349
 Db 308 NDAPVTFRTSEAGALWMSFNSTGAGALTQ----- 337
 QY 350 DGFNIYSFR-----LNGGETLYFMQKQSLIPASDINGQAGGVFEENFTVPSNQI 403
 Db 338 -GTTTYAMHQQGNLMAKRLIFQD-NQIINLKQSVQAGSLFPRDNVYVTSNGST 395
 QY 404 WQAGIHVSNSYVWYKNGVEHRLSKIGKTLHVAQKGNKSSISVGDGVILLEQAD 463
 Db 396 WTGAGIVVDNGSVYKQYKNGKGNLHKIGEGITLVQGTGINBGKLVKGDGVILNQAD 455
 QY 464 DQGNKQAFSEIGLVSGRGTVQINDKQFTDKFYEGFGRGLDNGHSLTFKRIQNTDS 523
 Db 456 NKGQYQAFSSVNTASGRPTVVLTDEROVNPDTWSWGKGLTDVNGSLTFHQLKADYG 515
 QY 524 AMYNHNTQAANVTITGNESTVLPNNNTNKLQDY-----RKEIAYNEM----- 567
 Db 516 AVLAN-NVDRKITL-----LDYALPDRVNLNWSGKGTACNLK 557
 QY 568 -----FGE-DRNKH----- 576
 Db 558 YNRPYNTTQYETLKQSTYGFPPDQSNATWSEVHSGQDAQVLVADRENTAGTLFHQ 617
 QY 577 -NGRLNLTK-PTTEDRTLLSGGTLNKDITQTKGLFSGRPTPYAHYK-----KR 629
 Db 618 LKGNLNVNRLPBGVTGALWMDGAAIDISGFTQENRRLTGSHVHAHYTQSVADKLA 677
 QY 630 WSEMGIPQGEIWDMDMNRTEKAKENPOLKGSAAVSNVSIKGMV-SNNANAFEG 688
 Db 678 SGDSVLTQPTSPSQDEKMERSTFDRSLKNTDPGLGNATL---NTTIQADNSVILG 734
 QY 689 -VPENQ-QWITICTRSDWTGLTTCQKYLDTKVNLS-IPQIQNGSINL----- 734
 Db 735 DSRVFIQDNOGQTAFTLEGTSAI-----KADKSVFNGYVNLNDSGLINI 782
 QY 735 -----TDNATANKGLAKLNGNTLTN-----HSQTLNNATQIGNR 773
 Db 783 NDIIFNGCIQANNSTVNISSDAVLNGLTSTALNLKCANALASQSFVSDGPNISDA 842
 QY 774 LQDNG-----TATVDNANL-----GNVHLTD----- 795
 Db 843 LSLNRPDEVSHTLLPYVYDASNNLKGDRLANVPYMSLSNINNOVGTYTLGEGE 902
 QY 796 -SAQSLAKSHFSHQIGDKG-----TVTLNNAWTMPEDCTLNTLANSIT 844
 Db 903 LSPDLTLQMLYSLPRGRNINLKSGLNAPATVSTMDTQWOMNNSFAGNMMLNTIYV 962
 QY 845 LNSAVASASNNTPRRSLTETTPISAHRNTLTVNGKLSQGGTFQFSLSGYSDTL 904
 Db 963 FNGG-----TSP-----FTTLTDINDAVQSAFWRTDL--NKAADL 997
 QY 905 KLSNABG-DYILSVRTGKEPETLEQLT-----VESKQNPUS 943
 Db 998 VIKSKTADGDSIV-NELKQSNKOTLDIPUSABAPATDULPASRVVGSVDTPIL 1056
 QY 944 DKLKFTLENDVDAKRYLVKNDGEFRLNPIKQELANDLVRAEQAEFTL-EAKQVE 1002
 Db 1057 SVAKEDGKKWVLDG--YQVARNQGGKAAATFMHISVNNFITVNNLNKRMEDRLDIN 1113

```

Qy 1003 PTAKTQGEKVSRRARAAPD-----TLPPQSLNALZAKQABETATQ 1049
Db 1114 GEAGTWY--RLINGSGSADGFTDHTTILOMGADRKHELSMDLFTGVMAITYTDTDSAD 1171
Qy 1050 KSKAXEKKQVSRKRAVSPDPLDOSLFALNALEVIADPOGE-KDRIAQEAEKOROKD 1108
Db 1172 LYSKTKS-----WGGGYASGLFPRSGAYFDVIAXYIHENKXDILNPAAGAKONFR-- 1222
Qy 1109 LISRYNSALSLSATVNSMLSVODELRLFYDOQASAVTNTIAQDKRXYDSDAFRAYQ 1168
Db 1223 -----SHSLYAGAEWGYRYHLT-----DTTFVEPQAEIWMGRL-----QG 1257
Qy 1169 OXTNLARQIGVOKALAN-----GRIGAVFSHSRSDNTFDOVKNHATLTMASGPAQYOW 1221
Db 1258 QTFHWNDSGMDVSMRBNVNPVGRGTGV-----SGKTFSGK--DMSLTARAGI-HYEF 1308
Qy 1222 GDLOFGVNVGTGISASKMAEEOQSKIRKAINGVNASYOPRIG--OLGIOPYFGVURY 1278
Db 1309 -DLTDSADVHLKDPAAGE-HQINGKDSKMLYGVGLNA-----RFGDNTRLGLE----- 1354
Qy 1279 FIERENTQSEERAVKTPSLAFNRY-----NAGIRVDYTF 1312
Db 1355 -VERS-----AFGKNTDIDAINNIR--YSF 1377

```

Search completed: June 23, 2004, 17:32:38
 Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2004, 17:29:18 ; Search time 56 Seconds
(without alignments)
7854.150 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273
Sequence: 1 MKKVFRINFLTACISLGIIV.....SQGSQLGKQGVKLGIRW 1394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_ricket:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5807	79.8	1436	2 08GQP3	08GQP3 haemophilus
2	5297.5	72.8	1391	2 08GQP2	08GQP2 haemophilus
3	4178	57.4	1449	2 09F325	09F325 neisseria m
4	4178	57.4	1449	16 09JMB4	09JMB4 neisseria m
5	4091.5	56.3	1444	2 07W111	07W111 neisseria m
6	4085	56.2	1457	16 09JXL6	09JXL6 neisseria m
7	4084	56.2	1457	2 08GKS4	08GKS4 neisseria m
8	4072	56.0	1457	2 03X7H1	03X7H1 neisseria m
9	4067.5	55.9	1454	2 09F324	09F324 neisseria m
10	2523	34.7	507	2 08KRT9	08KRT9 haemophilus
11	2128	29.3	505	2 08KGS5	08KGS5 haemophilus
12	2087.5	28.7	510	2 08KSC0	08KSC0 haemophilus
13	1839.5	25.3	526	2 08KSB9	08KSB9 haemophilus
14	1724.5	23.7	514	2 08KSN0	08KSN0 haemophilus
15	1676	23.0	1431	16 09JYX3	09JYX3 neisseria m
16	1654.5	22.7	1561	2 051169	051169 neisseria m

17	1641	22.6	1552	2 09A106	09A106 neisseria m
18	1620	22.3	523	2 08XSC1	08XSC1 haemophilus
19	1620	22.3	523	2 08KIE7	08KIE7 haemophilus
20	1544.5	21.2	185	16 09KCB4	09KCB4 neisseria m
21	1538	21.1	1764	2 09T134	09T134 haemophilus
22	1536.5	21.1	1773	16 09JVB9	09JVB9 neisseria m
23	1432	19.7	997	2 09S6X2	09S6X2 neisseria m
24	1431	19.7	993	2 09S6X3	09S6X3 neisseria m
25	1421	19.5	996	2 057309	057309 neisseria m
26	1417.5	19.5	996	2 057309	057309 neisseria m
27	1412.5	19.4	996	2 030574	030574 neisseria m
28	1397.5	19.2	992	2 030573	030573 neisseria m
29	1397.5	19.2	992	2 09S6X4	09S6X4 neisseria m
30	1393.5	19.2	992	2 09S6X5	09S6X5 neisseria m
31	1393.5	19.2	992	2 057035	057035 neisseria m
32	1113.5	15.3	1377	2 047692	047692 escherichia
33	1109.5	15.3	1377	2 088093	088093 escherichia
34	1089	15.0	1371	16 08CWC7	08CWC7 escherichia
35	1085.5	14.9	1372	16 083079	083079 shigella fl
36	1083.5	14.9	1372	2 054151	054151 shigella fl
37	1083.5	14.9	1373	2 09A158	09A158 shigella fl
38	1060	14.6	1376	2 083W66	083W66 escherichia
39	1051	14.5	1376	16 08FRM0	08FRM0 escherichia
40	1023.5	14.1	1335	2 091A54	091A54 escherichia
41	1015.5	14.0	1335	2 091A58	091A58 escherichia
42	935	12.9	1364	2 08VSU2	08VSU2 shigella fl
43	935	12.9	1366	2 054165	054165 shigella fl
44	931	12.8	1364	2 099QC6	099QC6 shigella fl
45	924.5	12.7	1364	2 084GK0	084GK0 escherichia

ALIGNMENTS

RESULT 1

08GQP3 PRELIMINARY; PRT; 1436 AA.

AC 08GQP3: 31-MAR-2003 (TRENBLER, 23, Created)

DT 31-MAR-2003 (TRENBLER, 23, Last sequence update)

DT 31-OCT-2003 (TRENBLER, 25, Last annotation update)

DE Adhesin (Fragment).

GN HAP.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

CC

OK NCBI_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=P860295;

RX MERLIN=22242173; PubMed=1235562;

RA Cutler D., Mason K.W., Howell A.P., Fink D.L., Green B.A.,

RA St Geme U.W. III;

RT "Immunization with Haemophilus influenzae Hap Adhesin Protects against

RT Nasopharyngeal Colonization in Experimental Mice."

RL J. Infect. Dis. 186:1115-1121(2002).

DR EMBL: A936379; AAN7923.1; -

DR GO: GO:0004295; P-tyrosin activity; IEA.

DR GO: GO:0006508; P-proteolysis and peptidolysis; IEA.

DR InterPro: IPR006315; Autotransport.

DR InterPro: IPR005446; Autotransporter.

DR InterPro: IPR001254; Peptidase_S1.

DR InterPro: IPR000710; Peptidase_S6.

DR InterPro: IPR004899; Peptidase.

DR Pfam: PF03797; Autotransporter; 1.

DR Pfam: PF02395; IGH1; 1.

DR Pfam: PF03212; Peptidase; 1.

DR PRINTS: PRO0921; IGASERPTASE.

DR TIGRFAMs: TIGR01414; autotrans_barl; 1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.

FT NON TER 1436 1436

SQ SEQUENCE 1436 AA: 159577 MW: C8B350983BDB6CB CRC64;

Query Match 79.8%; Score 5807; DB 2; Length 1436;
 Best Local Similarity 79.3%; Pred. No. 2, 1e-242;
 Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;

QY 1 MKKTIVRLNFLTACISLIGVSOAMAGHTYFGIDVYTRDPAENKGFYGAQNTIKYKXQ 60
 DB 1 MKKTIVRLNFLTACISLIGVSOAMAGHTYFGIDVYTRDPAENKGFYGAQNTIKYKXQ 60
 QY 61 GQVYVSTKAKMIDFVSVBNGVVALYVNOYTVSVANNGYTDVDFGABGNNDQHRFT 120
 DB 61 GQVYVSTKAKMIDFVSVBNGVVALYVNOYTVSVANNGYTDVDFGABGNNDQHRFT 120
 QY 121 YKIVKNNYKXONLHPEDDYHNPRLHKFVTEAPRIMTSMNGSYSPDRTKYERPRIG 180
 DB 121 YKIVKNNYKXONLHPEDDYHNPRLHKFVTEAPRIMTSMNGSYSPDRTKYERPRIG 180
 QY 181 SGRQFPRAND-----KGDVAGAYHYLTACNTINORAGANGSYLGQDVAKAGEYGLP 235
 DB 181 TGEQWMTDEBQKQSSKSMILADAYLMRLAGNTISQSGAGNGTVNLISGDTKPKNNYGLP 240
 QY 236 TASSKGBSGPMFTYDAKOKMLINGLMEBPPPEKENGFOLYRSYF-DEIFERDLHT 294
 DB 241 TGVSFGSSPMFTYDAKOKMLINGLMEBPPPEKENGFOLYRSYF-DEIFERDLHT 300
 QY 295 SLVTRAGNGYTTISGNDNGQSSITQ--KSGIPSEIKITLANKSLPLKEDKYNPRY-- 349
 DB 301 TELPRBNNGHTFTSNNGTGTVTQNEKVSHPQFKRTVQLNEMLKXK--EPYVAA 358
 QY 350 DGPNTVSPRLNNGEYTFMDQKQSLIFASDINGAGGLYFESNFTVSPNSNQTWQAGI 409
 DB 359 GGVVAYXPRLNNGKXITFGDRGTGTLTENNINQAGGLYFESNFTVSPNSNNAIMQAGAV 418
 QY 410 NVSNSTVYVWKNVGEHRLSKTGKTLHVOKSGNNGSSVGDGVYLLDQADQENKQ 465
 DB 419 HVGSDSTVWKNVGEHRLSKTGKTLHVOKSGNNGSSVGDGVYLLDQADQENKQ 478
 QY 470 AFSEIGVSGRGVQVQANDKQFDTKFEYFGFGGLDINGSLTFKRIQNTDEGAMTVN 529
 DB 479 AFKEVIGVSGRGVQVQANDKQFDTKFEYFGFGGLDINGSLTFKRIQNTDEGAMTVN 538
 QY 530 NTFQANVTITGNESTVLP-NGNNINKLYRRELYANNGPGEFTDKKNGKRNILYVPT 568
 DB 539 NTFQANVTITGNESTVLP-NGNNINKLYRRELYANNGPGEFTDKKNGKRNILYVPT 598
 QY 589 EDRLLLSGGTINLKGITQTKGLFESGRPTGHAYNHLNKRKSMGSIQGGIYWDHMI 648
 DB 599 EDRLLLSGGTINLKGITQTKGLFESGRPTGHAYNHLNKRKSMGSIQGGIYWDHMI 656
 QY 649 NRTKAPNFOIKGSAVVSRSVSSIEGNTVSNANNAATFGVVPQONTICTESDWTGLTT 708
 DB 657 TRTKAPNFOIKGSAVVSRSVSSIEGNTVSNANNAATFGVVPQONTICTESDWTGLTT 716
 QY 709 CQKVDLDIVYINSIPKQVINGSLTNDNATVANGKLAKNGVTLTNHSOTLSNNAQT 768
 DB 717 CKVDLDIVYINSIPKQVINGSLTNDNATVANGKLAKNGVTLTNHSOTLSNNAQT 776
 QY 769 IGNRRLSDNSTATVDNANLNGVHLTDSAQFSKSHSHOIQDGGITVTEJNAVTMP 828
 DB 777 TGNIOJSHNANATVDNANLNGVHLTDSAQFSKSHSHOIQDGGITVTEJNAVTMP 836
 QY 828 SOTLLQMLTLNNTITLNSAYSASSNNTTP-RRSLTETPTTSAHERFNTLTVNKLSCQ 887
 DB 837 SOTLLQMLTLNNTITLNSAYSASSNNTTP-RRSLTETPTTSAHERFNTLTVNKLSCQ 896
 QY 888 GTFQFTSSLFQYSDKAKISNDAGCYTILSVRTGKBEFTLEQTLTVESKDQNPISDYLK 947
 DB 897 GTFQFTSSLFQYSDKAKISNDAGCYTILSVRTGKBEFTLEQTLTVESKDQNPISDYLK 956
 QY 948 FTLENDVADGALRYLVYNDQFRLNMPKKEOELNDLVYABOERTEJAOVPEPTAKT 1007
 DB 957 FTLENDVADGALRYLVYNDQFRLNMPKKEOELNDLVYABOERTEJAOVPEPTAKT 1016
 QY 1008 QTGEKYSRBRARAFAFDTLTLDQSLNLALEKQKELTAFQSKAKTKYVSKRAVPSD 1067

DB 1017 QSNAPRVRK---RAVESDTPDQSQDVLQVQVYPTAEKQKKA--KRVSKRAVPSD 1071
 QY 1068 PVLDS-----LFALEALVETD 1085
 DB 1072 TLDDQSQDVLQVQVYPTAEKQKKA---KRVSKRAVPSD 1131
 QY 1086 APOQSEORLALAEAKRQKQKDLISHSNSALSELATVNSMLSYQDELDRLFYDQMS 1145
 DB 1132 APOQSEORLALAEAKRQKQKDLISHSNSALSELATVNSMLSYQDELDRLFYDQMS 1188
 QY 1146 AVNTIMADKRRYDSAPFRAYQOQKTLRQIVQKALNGHIGAVFSHRSNDTPDEQV 1205
 DB 1189 AVNTIMADKRRYDSAPFRAYQOQKTLRQIVQKALNGHIGAVFSHRSNDTPDEQV 1247
 QY 1206 NHATITMMSGFAQYQMGDLOFQVNVGTGISASKMAEQRXIKHKAIVGVNASYQFRLG 1307
 DB 1248 NHATITMMSGFAQYQMGDLOFQVNVGTGISASKMAEQRXIKHKAIVGVNASYQFRLG 1307
 QY 1266 QGIGQFQVNRYPFERENYQSEEVYKTPSLAFNRVYAGIRVDYFTPTDNI SVKPYFF 1325
 DB 1308 QGIGQFQVNRYPFERENYQSEEVYKTPSLAFNRVYAGIRVDYFTPTDNI SVKPYFF 1367
 QY 1326 VNVVDVSNANVQTTVALTVLQQFQRYWQKEVGLKALHFFQISAFISKSQSQSLQKQON 1385
 DB 1368 VNVVDVSNANVQTTVALTVLQQFQRYWQKEVGLKALHFFQISAFISKSQSQSLQKQON 1427
 QY 1386 VGVKLGYN 1394
 DB 1428 VGVKLGYN 1436

RESULT 2

Q8GQP2 AC Q8GQP2 PRELIMINARY; PRT; 1391 AA.
 AC Q8GQP2; 01-MAR-2003 (TREMELREL. 23, Created)
 DT 01-MAR-2003 (TREMELREL. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELREL. 25, Last annotation update)
 DE Adhesin (Fragment).
 GN Hap.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM K.A.
 RC STRAIN=TN106;
 RC MEDLINE=2242173; PubMed=12355362;
 RA Cutter D., Mason K.W., Howell A.P., Fink D.L., Green B.A.,
 St Gene J.W. III;
 RT "Immunization with Haemophilus influenzae Hap Adhesin Protects against
 RT Nasopharyngeal Colonization in Experimental Mice."
 RL J. Infect. Dis. 186:1115-1121(2002).
 DR EMBL; AF639380; AAN37924.1; -
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR006315; Autocatalysis and proteolysis.
 DR InterPro; IPR005546; Autocatalysis and proteolysis.
 DR InterPro; IPR009203; Cys_Ser_trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR000710; Peptidase_S6.
 DR InterPro; IPR004899; Peptidase.
 DR Pfam; PF02395; IGAL1.1.
 DR Pfam; PF03212; Pertactin.1.
 DR PRINTS; PR00921; IGASERPTASE.
 DR TIGRfam; TIGR01414; autocrans_Bact1.1.
 DR PROSITE; PS00135; TRYPSIN_SER.1.
 FT NON TER 1391 1391
 SQ SEQUENCE 1391 AA; A55385D8BC156FD1 CRC64;

Query Match

72.8%; Score 5297.5; DB 2; Length 1391;

Best Local Similarity 74.7%; Pred. No. 2e-220;
Matches 1052; Conservative 109; Mismatches 216; Indels 31; Gaps 15;

```

QY 1 MKKTVFRLNLTACISLISVSOAMAGHTYSDIYQYRPANRKGKFTVQAKQIKYVQ 60
DB 1 MKKTVFRLNLTACISLISVSOAMAGHTYSDIYQYRPANRKGKFTVQAKQIKYVQ 60
QY 61 GOLVGTSMTKAMIDSVSRNGVAALVENOYIVSAHVAVGYTVDVDEGAEANNPDQHRT 120
DB 61 GEMIGTMMKGVMPDLSMWVKGYSYTLISGHLISVAHVAVGYTVDVDEGAEANNPDQHRT 120
QY 121 YKIVRNNYKKNLHPEYEDDYNPLAKFTVTEAPLDQMTSNNGSTSDTKTPEYVIG 180
DB 121 YKIVRNNYKKNLHPEYEDDYNPLAKFTVTEAPLDQMTSNNGSTSDTKTPEYVIG 180
QY 181 SGQRQVNDQK--GDVGAAYHLLTAGTNNRQGANVAYLGGDVAKAGEPLTIG 238
DB 179 SGQVMMKDKNNKTIQDLAYG--SMILGNTDEDPAGNGLLELNGRVQMPKGTPLTIG 237
QY 239 SKDGSQPMETVDAEKQWLLINGILREGNPEPEKENGQPVYKSYDEIEERDLHTSLYT 298
DB 238 SPEDSSPMETVDEKQWLLINGILREGNPEPEKENGQPVYKSYDEIEERDLHTSLYT 297
QY 299 RAGNGVYIISGNDG-----QSSITQKSGIFSEIKITLANNGL-PLKXQKVNPRYDGP 352
DB 298 TNAEYKFNIGSDHNGVATIKSTLPKKAIOBERVSLYDNLHARDKNDSPSYKGP 357
QY 353 NIYSPLANNGETLVFMOKQSLIFASDINAGAGLYPFGNFTVSPNSQ--TWQAGAHV 411
DB 358 NPSPALHNGKSIYFGQGTGTLTENNINQAGAGLTFBGRNRYVGNMNTWQAGAVY 417
QY 412 SENSTVWKVNGVENHRLSKIGKGLVQAKNGKSGISVQKXVILEQDADQGNQAF 471
DB 418 GESSTVQVYHNEPEGLSKIGKGLVQAKNGKSGISVQKXVILEQDADQGNQAF 477
QY 472 SRIGLVSGRGTYQINDKQPTDKRYFGFRGGSLDLNGSLTFPKAIQNTDEGAMTYNAT 531
DB 478 KEVGIYSGRATVQINSADQVDPNNIYFGFRGGSLDLNGSLTFPKAIQNTDEGAMTYNHA 537
QY 532 TOAANTTNGESITLVNNGNINKLQYREIYANNGGEGDNNKNGNGLNIYKPTEDR 591
DB 538 SOTANTTNGNNTI---NSDSKQNLNKKDLAFNMGEGEDDRAKNGNGLNNTQCPNAAEN 593
QY 592 TLLSGGTLNKGDIYOTRGKLPFGSPPTPHAVNHLNRRWMBEGIPQGIYWDHDMINT 651
DB 594 HLLSGGTLNNGNITLQNGGLTVFSRPPPHAVNHLRDLDSNMEGIPQGIYWDHDMINT 653
QY 652 PKAENFOIKGSAYVSRVSSISGKWTYSNNAATPGVYNQANTICTRSWTGLTTCOK 711
DB 654 PKAENFOIKGSAYVSRVSSISGKWTYSNNAATPGVYNQANTICTRSWTGLTTCOK 713
QY 712 VDLTDKRVINSIPKQIUNGSLNLTDAATANKGLAKNGAVTLNHSOFTLSNNAQIQA 771
DB 714 VDLTDKRVINSIPKQIUNGSLNLTDAATANKGLAKNGAVTLNHSOFTLSNNAQIQA 773
QY 772 IRLNSTAYVDNANLNGVNLTDLSAQPSLNKSHFSHQIGDKTIVTLNENATVWPSDT 831
DB 774 IKLSHNAATVDNANLNGVNLTDLSAQPSLNKSHFSHQIGDKTIVTLNENATVWPSDT 833
QY 832 TLQNLINNSTITLNSAYSSNNT--RRSLETTTPPSAHRPNTLTYNGLSQGG 889
DB 834 TLQNLINNSTITLNSAYSSNNT--RRSLETTTPPSAHRPNTLTYNGLSQGG 893
QY 890 FQFTSLFGYKSDKLTLSNAEDYILSVRTGKEPTELLQTLVESKDNQPLSDKLEF 949
DB 894 FQFTSLFGYKSDKLTLSNAEDYILSVRTGKEPTELLQTLVESKDNQPLSDKLEF 953
QY 950 LENDVDAGALRYKLVKNDGFRILNPIKEQHLNDIYRAEQERTLAKQYETACTQT 1009
DB 954 LENDVDAGALRYKLVKNDGFRILNPIKEQHLNDIYRAEQERTLAKQYETACTQT 1013
QY 1010 GERKVSRRAAAPPTLDDQSLNALEAKQALTEPQKSAKTKYVSKAV--FSQ 1067
DB 1010 GERKVSRRAAAPPTLDDQSLNALEAKQALTEPQKSAKTKYVSKAV--FSQ 1067

```

```

DB 1014 SKAVRSR---RAVSDPLPQSLKALPAKQ-LITTEQTSQA--KKVRSKARAREFS 1267
QY 1068 PLIDQSLFALEALVDAIPQOSEKRLAQEAEK-QRKQKDLISRYNSALSELATV 1126
DB 1068 TLDPQ---LIGALVIDAQQVKKPQTEEBEERQKKEILISRYNSALSELATV 1124
QY 1127 SMLSVQDELDRLFTVQASAVNTNIAQDKRYDSAPRAVQOQKTLRLQIQKALNGR 1186
DB 1125 SMLSVQDELDRLFTVQASAVNTNIAQDKRYDSAPRAVQOQKTLRLQIQKALNGR 1183
QY 1187 IGAVFSHSRNDTPDEQVQNHATLMMSGFACQMGDIQGVAVNGTIGIASGAEGSRK 1246
DB 1184 IGAVFSHSRNDTPDEQVQNHATLMMSGFACQMGDIQGVAVNGTIGIASGAEGSRK 1243
QY 1247 IHRKAIYGVNAYSQFRLGQIGIOPYGVNRYFIERENYQSEBEVRYTESLAFNRNNGI 1306
DB 1244 IHRKAIYGVNAYSQFRLGQIGIOPYGVNRYFIERENYQSEBEVRYTESLAFNRNNGI 1303
QY 1307 RVDYTFPTNTNISVRYFFNVYDVSNVQTTNLTVLQQPFGRITWQXVGLKAEILHF 1366
DB 1304 RVDYTFPTNTNISVRYFFNVYDVSNVQTTNLTVLQQPFGRITWQXVGLKAEILHF 1363
QY 1367 QISAFISKQSQSLQKQNVGKLGVRK 1394
DB 1364 QISAFISKQSQSLQKQNVGKLGVRK 1391

```

RESULT 3

Q9FJ25 PRELIMINARY; PRT: 1449 AA.

AC Q9FJ25, 01-MAR-2001 (Tremblrel, 16, Created)

DT 01-MAR-2001 (Tremblrel, 16, Last sequence update)

DI 01-OCT-2003 (Tremblrel, 25, Last annotation update)

DE Adhesion penetration protein.

GN APP.

OS *Neisseria meningitidis*.

OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales.

CC Neisseriaceae; Neisseria.

OX NCBI_TaxID=487;

RY [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Z3515;

RA Ala-Aldeen D.A.;

RT Autotransporter Protein of *Neisseria meningitidis*;

RT *Identification and Characterisation of App: an Immunogenic Autotransporter Protein of *Neisseria meningitidis*;

RL Submitted (0CT-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ296276; CAC14670.1; -

DR MEROPS: S06.006; -

DR GO: GO:0008233; F:peptidase activity; IEA.

DR GO: GO:0004295; F:trypsin activity; IEA.

DR GO: GO:0005508; P:proteolysis and peptidolysis; IEA.

DR InterPro: IPR006315; Autolysinport.

DR InterPro: IPR005546; Autolysinport.

DR InterPro: IPR009603; Cys Ser trypsin.

DR InterPro: IPR001254; Peptidase_S1.

DR InterPro: IPR00710; Peptidase_S6.

DR InterPro: IPR004899; Peptidase.

DR Pfam: PF02395; IGAL1; 1.

DR Pfam: PF03212; Peptidase; 1.

DR PRINTS: PR00921; IGASRPASE.

DR SMART: SM00020; TRYP_SPC; 1.

DR TIGRFAMs: TIGR01414; autotransp_bar1; 1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Protease; Serine protease.

SEQUENCE 1449 AA; 159071 MW; C07850BDAB47C25D CRC64;


```

17 YLALCSFGLPQAAAGHTYGTNGYDPEHNGKGFANAGADIVYVKKKGELVGSMT 76
QY 70 KAPMIDPSVSSNGVAALVENQYIVSAHVAVGTYDDFGABGNPDORHFTYKVKXNY 129
Db 77 KAPMIDPSVSSNGVAALVQDYIVSAHVAVGTYDDFGABGNPDORHFTYKVKXNY 136
QY 130 KK-DNHAPEDDYHNPRLHFTYEAAPIMTNNMGSTYDRIKHEPRTIGSGRQWRN 188
Db 137 KAGTNGPYGGDYHNPRLHFTYDAPVEMISTYMDGRYITDNNYDREYRGAQRQWRN 196
QY 189 DQDKGD-----QVAGAYHYLTAGNTNORAGANGSYLGGDVKXAGEPLIAGSKGD 242
Db 197 DEDEPNRRBSYHIAISALVGCNTFAQNSGGGTVALSSEKIKMSPGFLPGSGFRD 256
QY 243 SSGEMPIYDAEQKWLINGILRBGNPBGKNGPOLYKSYR-DEIFERDHTSLYTRAG 301
Db 257 SSGEMPIYDAEQKWLINGVLQTNPNPYTGKNGPOLYKSYRDEIFAGDTHSYFPRQ 316
QY 302 NGVYTTISGNDNGQSITQK---SGIASEIK---ITLNMSLPLKEDKXHNPRDGNIT 355
Db 317 NGKTSFHDNNNGTGKIDAKHEHNSLPRLKTTVQLFVNSLSERDAEPYTHAA-OGVNST 375
QY 356 SPRLNNGETLYFMDQKQSLIFASDINOAGGLYFEGNFTVSPNSQTOAGAGIHWENS 415
Db 376 RPRLNNGENISPIDGKGLILTSININGAGGLYFQDPFTVSPENNETWQAGVHISDS 435
QY 416 TVTKVNGVHBRSLKLGKTLVQANGENKSTIVGKVIIEQADQDQKQKAFSEIG 475
Db 436 TVTKVNGVANGBRSLKLGKTLVQANGENKSTIVGKVIIEQADQDQKQKAFSEIG 495
QY 476 LVSRKGTVOLNDDKQDPTDKFYFGRGRGLDNGSLFKRIQNTDEGATVNHNTQAA 535
Db 496 LVSRKGTVOLNDDKQDPTDKFYFGRGRGLDNGSLFKRIQNTDEGATVNHNTQAA 555
QY 536 NVITIGNESIVLPNGNINIKLDYRKEIAYNGFGBTDKRNKNGELMIYKPTTEDRLLL 595
Db 556 TVTITGKQDIA-TTGNN-NSLDSKKEIAYNGFGBTDKRNKNGELMIYKPTTEDRLLL 613
QY 596 SGGTNLAKGDIYOTKGLFSSGRPTPRPANHNLKRNSEMEGIPQELIYWHDMINRTFAR 655
Db 614 SGGTNLAKGDIYOTKGLFSSGRPTPRPANHNLKRNSEMEGIPQELIYWHDMINRTFAR 673
QY 656 NFOIKGSAVSNVSNSEIEGMMVTASNANATFEGVVPNOQNTICRSRDTGLTQCKXDLT 715
Db 674 NFOIKGSAVSNVSNSEIEGMMVTASNANATFEGVVPNOQNTICRSRDTGLTQCKXDLT 733
QY 716 DTKYINSIPKTOINGSINTMATANVGLAKLNGVTLNHSQFTISNNAFO-----I 769
Db 734 DTKYIASITWTDISGNVSLADHAIHLNLTALTNLNSLNGDTRYVSHNATQNDLSLV 793
QY 770 GN-----IRLSDNSTATVDMANKNHNHLTDS 796
Db 794 GNADATNQAATNGNTSASGNASFNISDHAOVNGSILTSNAGKANYSHSALNGVNSLADK 853
QY 797 ADFSLKNSHSHQIOGDKTTLTLENATWMPEDTTLONTLNNSTITLNSAY----- 849
Db 854 AVFHEBESRFTQISSSKOTALHLDSEWLTSPGTELCNMLDMLATITLNSVRRDAGA 913
QY 850 -SASSNTPR---RSLTETPTPSABHRNTLTNGKLSQGTFOFTSLFGYSKDLK 905
Db 914 QGSAATDAPRRRRESLSVTPPASASHNTLTNGKLNQGTFRFMSLEFGYSKDLK 973
QY 906 LSNABGDYILSVRNGKPELTGLVYSKONP.SDKLKTLENDVDVAGARVYLV 965
Db 974 LASSBEGTYTLAANNNGHBPASLBOLTVVGKUNP.SENINFLTLQNEHDDKAKRYLYI 1033
QY 965 KMDGFIHNPPIKQELHNDLVRAOERTLSAKOVBETATYOTGEKYSRRARAAAP 1025
Db 1034 KQDEFFRHNVPVKEGLSKGKAE-----AKQKGNKQNSIDLIALIAGR 1079
QY 1026 DTLPDQSLNALFAKAE-----LTAFETQSKAKTKYVRSKRAVSDPLTQOSIFALFA 1079

```

```

Db 1080 DAV-EKTESVAEPARQAGENVGIMQAEFEKKRVQADK-----DTALAKQREAEETP 1130
QY 1080 AAEVVDAPQOSEMDRLAQSEAKO-RKQNDLISRYNSNALSATNSVLSVODELRL 1138
Db 1131 ATTAPPRARRRRRD-LPQPPQPPQPPQRLISRYNSNALSSEFATLNSVFAVQDELRV 1189
QY 1139 FVDQKASVMTNIAQDKRRYDSDAFRAYOOQKNTLRLQIGVQKALANRIGAVSHSRSDN 1198
Db 1190 FAEEDRRNVAWTSIGRTKHYRQDFRATQO-TCAPQIQMOKNGSGRVALLESNRIEN 1248
QY 1199 TFDQVKNKATLTMWSGPAQYQWGLQPSVAVGTGISASNAEBSKIRKAKYGVNA 1258
Db 1249 TFDQVKNKATLTMWSGPAQYQWGLQPSVAVGTGISASNAEBSKIRKAKYGVNA 1308
QY 1259 SVQERLQGLQIPFPVNNRYTERENQSEEVKTPSLAFNRNAGIRVDYETPTNTI 1318
Db 1309 RRAQEGFGFIEPATRYFQKADRYENVNIAPGAEFRFRAGIKADYEFKADQHI 1368
QY 1319 SVKQYFPNVYVDSNANVOVTNLTVLQOPFGRYWQKEVGLKAEILHFQLSAFISKQGS 1378
Db 1369 SITPILSISTYDASGKVRTRVNTAVTLADPQKTRBAEGVNAELKPTLISHAAAKGP 1428
QY 1379 QLGKQNVGVKGYRW 1394
Db 1429 QLEAQHSAGIKLGYRW 1444

RESULT 6
Q9UXL6 PRELIMINARY; PRT; 1457 AA.
ID Q9UXL6
AC Q9UXL6;
DT 01-OCT-2000 (TREXBLrel. 15, Created)
DT 01-OCT-2000 (TREXBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREXBLrel. 25, Last annotation update)
DE Adhesion and penetration protein.
GN NMB1395.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B.
RX MEDLINE=20115755; PubMed=10710307;
RA Tettelein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Felsen J.A., Kechoud K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Grim M.L., Deboy R., Peterson J.D., Hickey E.K.,
RA Hatt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parkey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Ueberback T.R., Kouri H., Qin H., Vamathavan C.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.W., Moxon E.R., Rappaport R., Venier J.C.;
RT Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.
RL Science 287:1809-1815 (2000).
DR EMBL; AE002547; AAF42312.1; -.
DR PIR; D81019; D81019.
DR MEROPS; S06_006; -.
DR TIGR; NMB1395; -.
DR GO; GO:0004295; "Zinc peptidase activity"; IEA.
DR GO; GO:0006508; "Protein catabolic process"; IEA.
DR InterPro; IPR006315; Autocatalytic domain.
DR InterPro; IPR005546; Autocatalytic domain.
DR InterPro; IPR001254; Peptidase S6.
DR InterPro; IPR000710; Peptidase S6.
DR InterPro; IPR004839; Peptidase S6.
DR Pfam; PF02395; IGA1; 1.
DR Pfam; PF03212; Peptidase S6.
DR PRINTS; PRO0921; IGASERPTASE.
DR SMART; SM00020; TYP_SPC; 1.
DR TIGRfam; TIGR01414; autotrans bar1; 1.
DR PROSITE; PS00135; TRYPIN_SER; 1.

```

Db	Accession	Protein Name	Length
Db	994	KLMALASSEGGHYTLAVNNTGNBPASLEQLTVEGSKONKPLESEMLNFTLQNEHVDAGAMRY	1043
Qy	963	KLVKNDGGERLHNPPIKEQGLANDVARE--QAEK-----TLKAKVEPTACTQT-GE	1011
Db	1044	QLIRKKGCEFRLLHNPVVEQELSPDKLGNKAKKQAEKDNANOSIDALLIAQRDAVEKTSVAE	1103
Qy	1012	PKYRSRRRAAPAFDILLPDQSLTNALAAKQAELETFQTSKATKTKVRSKRAVSPDLL	1071
Db	1104	P-----RQAGGEV--GIMQSEEEK-RVQNDKDTALAKRALETPATTAFP--	1149
Qy	1072	QSLFALAEALVIDIADPQSEKDRLOAEAEKQKQKODLISYNSNLSLSEATYNSMLAV	1131
Db	1150	-----RARRRRRLDQLQDPQGPQGPQGRDLISRYNSGLSEFSLTNSVFAY	1192
Qy	1132	QDELDELFDYDAQSAVYTNIAQDKRIDSPAPRAYQQQKTNLQIQVAKALANRIGANF	1191
Db	1196	QDELDFVAEERRRRAVWTSGIDPTVYRSQDPFARQQ--TDLQIQMCKNLSGRVGLLE	1254
Qy	1192	SHRSRNTEFDEQVYNNHLLTMMSGPAQYQWGLQGLQGVNVGGSASAKMAEQSKIRKXA	1254
Db	1255	SHRNTNTEFDGIGNSRLAHGAVFGQIGIDRFYIGISAGAGSSGLSDGIGKIRRY	1314
Qy	1252	INVGVASAYCFRLGQLGIQIPYGVNRYFIERENYQSEEVRYKPPSLAFRRYNAGIRVDYT	1311
Db	1315	LHVGIGARRRAGFGGEGIEPIGHATRFVQKADYEMVNLAPGLAFNRVAGIKADYS	1374
Qy	1312	PPTDNISKRYEYFFVNTVDSNVANVQTTNLTLYLQDPGRYQKEVGLKAEILHFQISAF	1371
Db	1375	FXRQHSITTFYLSLSTYDAAGSKVRRTVNTVLAQDGKTRSAEGVNAIBKFTLSLH	1434
Qy	1372	ISKSGSGLGKQKQNVGKLGYYRW	1394
Db	1435	AAAHKGPQLAEQHSAGIKLGTRW	1437

RESULT 7

08GKS4 PRELIMINARY; PRT: 1457 AA.

ID 08GKS4

AC 08GKS4;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE App.

OK App.

OS Neisseria meningitidis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

OC Neisseriaceae; Neisseria.

OX NCBI_TaxID=487;

EN [1]

EN SEQUENCE FROM N.A.

RP STRAIN=H44/76;

RC STRAIN=H44/76;

EA van Ulsen P.;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL, AY150285; AA01716.1; -

DR GO; GO:0004295; P:trypsin activity; IEA.

DR GO; GO:0006058; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR006315; Autotransport.

DR InterPro; IPR005546; Autotransporter.

DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR000710; Peptidase_S6.

DR InterPro; IPR004899; Pertactin.

DR Pfam; PF03797; Autotransporter; 1.

DR Pfam; PF02395; IGAL; 1.

DR Pfam; PF03212; Pertactin; 1.

PRINTS: PR00921; IGASERPTASE.
 DR TIGREPAMS; TIGR01414; autocorans bar1; 1.
 DR PROSITE; PS00135; TRYPSIN_SER_1.
 SQ SEQUENCE 1457 AA; 159965 MW; B579F52A5CDD768F CRC64;

Query Match 56.2%; Score 4084; Da 2; Length 1457;
 Best Local Similarity 55.2%; Pred. No. 4,8e-168;
 Matches 805; Conservative 213; Mismatches 340; Indels 100; Gaps 16;

```

QY 10 FLNCLSLGVSQAAAGHTYFDYQYTPDPAENKGTGTAQNTKYNKQQLVGSMT 69
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 27 YLAICISFGLIPQAAAGHTYFGLNYQYDFPENKGFAGACDIYVYKKGELVGSMT 86
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 70 KAPMIDSVSSNGVAALVENQYVSAHVAVCYTDVDFGAEGNPQHFYTKYKKNY 129
DB KAPMIDSVSSNGVAALVGDQYVSAHVNGYVNDVFGAGNRPQHFYTKYKKNY 146
QY 87 KAPMIDSVSSNGVAALVGDQYVSAHVNGYVNDVFGAGNRPQHFYTKYKKNY 146
DB KAPMIDSVSSNGVAALVGDQYVSAHVNGYVNDVFGAGNRPQHFYTKYKKNY 146
QY 130 KCDML-HPEYDDYHNPRLHKFVTEAPIDMTSNMGSTYSDRKYERVRVIGSGRPWRN 188
DB KAGTKAHRYGQDTHMPRLHKFVTEAPIDMTSNMGSTYSDRKYERVRVIGSGRPWRN 206
QY 147 KAGTKAHRYGQDTHMPRLHKFVTEAPIDMTSNMGSTYSDRKYERVRVIGSGRPWRN 206
DB KAGTKAHRYGQDTHMPRLHKFVTEAPIDMTSNMGSTYSDRKYERVRVIGSGRPWRN 206
QY 189 DQDQD-----QVAGAHYTLAENTHNRGANGCYSLGDVRAKGEVPLPLAGSKD 242
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 DEDPPNRRESSYHIASYMLVGNTFQMGSGAGTVMASBEKIKSPYGLPFGSSFGD 266
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 SGSPMFLYDAEKOKMILINGILREGNPESKENGAPOLVRSYF-DEIFEDILTSLYTRAG 301
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 SGSPMFLYDAEKOKMILINGILREGNPESKENGAPOLVRSYF-DEIFEDILTSLYTRAG 326
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 NGVYVTSIGNDGSGSTYK--SGIPBEIK--TILANMSPLEKOKVHNRYDGNIT 355
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 327 NGKYSFVDONNGTGKINKAHEHNSLPRLKRTVQLFVNSLSEIAREPYHA--GVVNSY 385
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 356 SPRLNERTLYFPCQKOSLIFASDINQAGGLYFEGNPTSPNSQNTQAGGIVHSENS 415
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 386 RPRLNENISFIDEKCELLITNSINQAGGLYFQSDPTVSPNNEITQAGGIVHSENS 445
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 416 TVYKVGNGVGHDLSTKIGKTLHVQAKENGSGISVSGKVIILEQDADQKQKQAFSHIG 475
DB TVYKVGNGVGHDLSTKIGKTLHVQAKENGSGISVSGKVIILEQDADQKQKQAFSHIG 505
QY 446 TVYKVGNGVGHDLSTKIGKTLHVQAKENGSGISVSGKVIILEQDADQKQKQAFSHIG 505
DB TVYKVGNGVGHDLSTKIGKTLHVQAKENGSGISVSGKVIILEQDADQKQKQAFSHIG 505
QY 476 IVSGRGTVALNDKQDFTDKFYFSGFRGRDLNGSHLTKRKIQNTDGBAMIVNNTQAA 535
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 506 IVSGRGTVALNDKQDFTDKFYFSGFRGRDLNGSHLTKRKIQNTDGBAMIVNNTQAA 565
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 536 NWTITGNESTVLPNGNINIKLDYRKELAVNGRGEDTKNGKNGIMLYKPTERTILL 595
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 566 TVTITGKQDIA-TTGNN-NSLDSKKEITAVNGFGEKDTTKTNGKLMVYQPAEDRTILL 623
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 596 SGGTNLKGADITQTKGLFFSGRPTRPHAYNHLNKRWSNEMEGIPQGLVWMDMINETFAE 655
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 624 SGGTNLKGADITQTKGLFFSGRPTRPHAYNHLNKRWSNEMEGIPQGLVWMDMINETFAE 683
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 656 NFQIKGSAAVSNVSSIBGMWTVSNANADTPGVVPNQNTTICRSIDWGLTTCQKDLT 715
DB NFQIKGSAAVSNVSSIBGMWTVSNANADTPGVVPNQNTTICRSIDWGLTTCQKDLT 743
QY 684 NFQIKGSAAVSNVSSIBGMWTVSNANADTPGVVPNQNTTICRSIDWGLTTCQKDLT 743
DB NFQIKGSAAVSNVSSIBGMWTVSNANADTPGVVPNQNTTICRSIDWGLTTCQKDLT 743
QY 716 DPKVNSLPTKQIKSINLTQNTAVYKGLAKINGVTLTHSOFILSNNAIQGN----- 77-
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 744 DPKVNSLPTKQIKSINLTQNTAVYKGLAKINGVTLTHSOFILSNNAIQGN----- 803
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 772 -----RLSDNSTATVNDNANLNTDS 796
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 804 GNAQATPQATLNGNTSASGNASFWLSDEAVNGSLLISGAKNAVSHSLNGVSLADK 863
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 797 AQSLSKSHFSHOIQDQKGTVYTLNATWYMBEDTLQNLTLNNTSTILNLSAY----- 849
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 864 AVFHESSRFTQGISGKQTAHLKDSMTLPJGTELGNLMDATLTNLSARDDAAGA 923
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 850 -SASGNTPR-----RSLTETPTSAERHNTLVNKGSLSGQGFQPTSLFGYKSD 902
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 924 QIGSATDAPRRSRRSRLSVLPPTVSERNTLVNKGSLSGQGFQPTSLFGYKSD 983
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 903 KKLSTNDAEGDYLLSVRNTGKEPETLRLQTLVSKQNPQSLDKLFTLENDHYDAGALRY 962
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 984 KKLSTNDAEGDYLLSVRNTGKEPETLRLQTLVSKQNPQSLDKLFTLENDHYDAGALRY 1043
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 963 KLVNDGEFFRLNLEIKQGLNDLVRAQMRLEKQVEPNTQYGERKYSRRARA 1022
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1044 QLRDGEFFRLNLEIKQGLNDLVRAQMRLEKQVEPNTQYGERKYSRRARA 1089
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1023 APPDTLPQOSLNLNLEKQAE-----LTAEQSKYAKTKVSRKAVFSDPLDQSLFA 1376
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1090 AGEDAV-EKISVAERPRQGENVGYQAEBEKRVQADK-----DTALAKREAE 1140
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1077 LBALEVIDAPQOSEKDRPLADEAKRQKQDLSISYNSALSLSLATSNSMLSVDELD 1136
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1141 TRPATTPARRARARBDLPQLOPQOPQOPQDLSISYNSALSLSLATSNSMLSVDELD 1200
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1137 RLFPDQASAVNTYADQKRDSDAPFAYQOQKTLRLQYQALANGSLGAVFHSNS 1196
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1201 RYFAEDRRNAAVMTSGIDTYHRSQDFRNVROO-TDLRQIGMOKLASGKVGILFESHNT 1259
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1197 DNTFEOYKNAATLTMWSPAYQYQWGLQSGVNVGTGISASQAEQSRKIHKKAINGV 1256
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1260 ENTFDQIGNSARLAKAVFQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 1319
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1257 NASYQFRLQGLQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 1316
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1317 NISVAFYFVNVYVDSANVQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 1376
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1380 HISTPEYLSYLTAAAGKRRVAVTAVLADQKTRSAQWNAEIKGFTLSLHAAAK 1439
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1377 GSQLOQKQNVGKLGYYW 1394
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1440 GPQLEKQHSAGIKKGYRW 1457
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 8

```

Q9X7H1 PRELIMINARY; PRT: 1457 AA.
AC Q9X7H1;
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE App. protein.
GN App.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-15:P1.16;
RA Abdel-Hadi H., Woolridge K.G., Ala Aldeen D.A.;
RT "Identification and characterization of neisserial App.":
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ242535; CAB43832.2; -.
DR MEROPS; S06.006; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR000710; Peptidase_S6.
DR InterPro; IPR004899; Peractin.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF03395; IGA1; 1.
DR Pfam; PF03212; Peractin; 1.
DR PRINTS; PR00921; IGASERPTASE.
DR SMART; SMO0020; TYP_SPC; 1.
DR TIGREPAMS; TIGR01414; autocorans bar1; 1.
DR PROSITE; PS00135; TRYPSIN_SER_1.

```

KW Hydrolyase; Protease; Serine protease.
 SQ SEQUENCE 1457 AA; 159957 MW; CE62390B6C04B781 CRC64;

Query Match 56.0%; Score 4072; DB 2; Length 1457;
 Best Local Similarity 55.0%; Pred. No. 1.6e-167;
 Matches 804; Conservative 219; Mismatches 330; Indels 110; Gaps 19;

```

10 PLTACISLGIIVQAMNGHTYFGIDYQYRPFANRKGKFPVQAKIKYVNGQOLVGTMT 69
27 YIATLSFGILPQAMNGHTYFGINQYRPFANRKGKFPVQAKIKYVNGQOLVGTMT 86
70 KAPMTDFSVSRNGVAALVGDQYIVSAHNGVANNVDFGABRPDDHRTTYKIVKNNY 129
87 KAPMTDFSVSRNGVAALVGDQYIVSAHNGVANNVDFGABRPDDHRTTYKIVKNNY 146
130 KQDNV-HPYEDDHNRLKFTYTAAPIDMTSMNGSTYSRKYPRKAVYISGGQFPRN 188
147 KAGTGHFGYGGDYHMPRLKFTYTAAPIDMTSMNGSTYSRKYPRKAVYISGGQFPRN 206
189 DDDKGD-----QVAGAYHYLTAGNTHQKQAGNGSYLGQVRAKAGEYPLPLIAGSKD 242
207 DEDERNRBSSTYHLSYSLVSGNTFPAQNSGGGTVMGSEKIKHSPYGLFTGSSPFD 266
243 SSGPMFTYDAEKQKMLINGILRBNPFEKNGPQLVKRSTF-DEIFRDLATSLYTRAG 301
267 SSGPMFTYDAEKQKMLINGILRBNPFEKNGPQLVKRSTF-DEIFRDLATSLYTRAG 326
302 NGYITISGNDGSSITQK---SGIPSEIK---ITLANNSLDKKKKKNNPYDGPNTY 355
327 NGKYSFNDNNSTKTKINAKENHNSLPNLRKRTYQLPNSVLSRTAREVYHAA-GGVNSY 385
356 SPLNNGETLYFMDQKQSLIPASDINOAGLTFYBGNFTVPSNNSQWQAGIHVENS 415
386 RPLNNGENISFTIDEGGELLISNINOAGGLYFGQDTPVPENNETWQAGVHSEDS 445
416 TYTWKNGVEHDLRSKIGKGLTHVQAKSENGKSTISVQDKVILLEDQADQKQAFSEIG 475
446 TYTWKNGVEHDLRSKIGKGLTHVQAKSENGKSTISVQDKVILLEDQADQKQAFSEIG 505
476 LVSGRGTQLNDQKQPTDQKFGYRGRGRDLNQHSHSTFRLQNTDDEAMVYNNHTQAA 535
506 LVSGRGTQLNDQKQPTDQKFGYRGRGRDLNQHSHSTFRLQNTDDEAMVYNNHTQAA 565
536 NVITIGNESIVLPNGNNINLQYRKEIANGMFGSETDKNKNGRNLILYKPTTEETLL 595
566 TVITIGNKDIA-TTGN-NSLDSKKEIANGMFGSETDKNKNGRNLILYKPTTEETLL 623
596 SSGTNLKGDIQYTKGLFFSGRPTPAVYNNHNRKMSMEGIPQGETVWDHDMNIRPKAE 655
624 SSGTNLNGNITQYTKGLFFSGRPTPAVYNNHNRKMSMEGIPQGETVWDHDMNIRPKAE 683
656 NFOIKGSAVSRVSSIEGNVTVSNANATFQYVPOQNTICTSDMTGLTQCKYDLT 715
684 NFOIKGSAVSRVSSIEGNVTVSNANATFQYVPOQNTICTSDMTGLTQCKYDLT 743
716 DTKVINSIKPTQINSIMLTNATAVNKGK----- 745
744 DDKVINSIKPTQINSIMLTNATAVNKGK----- 803
746 -----AKINGNVTLNHSQFTLSNNATQIGNRISNGSTAYTYDNANINGVHLLTDS 796
804 GNAQATFNQATLNGNTSASGNASFNLSHDAVNGSLTSLSNAAYNSHSLANGVSLADK 863
849 AVEHSSSRFTQISGGKDTALHLCQSEMTLPSGELGANLNNATITLNSAYRHNKAKA 923
850 SASNVNTPR-----RSLETETFTPSAHRPNTLYVNGKLSGGQFTQTSLSFGKSD 902
924 QTSATDAPARRRSRRSRSLSVTPPLVSSEFNTLYVNGKLSGGQFTQTSLSFGKSD 983
903 KTKLSNDARGYTILSVRNTGSEPTILEQTLVSDNQPSLQCKKFTLENDHNDAGLRY 962

```

```

DB 964 KTKLSNDSRGYTTAAVNTNGENPASILEQTLVYBGKDKKPLSENLTPLQNEHVADGANKY 1043
QY KLVNNDGEFFRLNHPKQDELNDLVRAE---QABR-----TLBAKQVEPTAQTQ-GE 1011
DB 1044 QLRKQGEFFRLNHPKQDELNDLVRAE---QABR-----TLBAKQVEPTAQTQ-GE 1103
QY 1012 PKVSRRAAPAAFPDTLPDQSLNLEAKQDELTAETQSKAKTKKPSKQAVSDPDL 1071
DB 1104 P-----ARQAGEHN---GIMQAGEEK---RVQADKDTALQREAFETRPATYAF- 1149
QY 1072 QSLFLEALAEVIDAPQOSEKDLQAEAKQKQKDLISYNSALSELATVNSMLV 1131
DB 1150 -----BARARLDLPQLQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 1195
QY 1132 QDELRLDFVQDQSAVMTNLAQDKRRYDSDAFRAYQDQKTNLQISVQPLANRIGAVF 1191
DB 1196 QDELRLDFVQDQSAVMTNLAQDKRRYDSDAFRAYQDQKTNLQISVQPLANRIGAVF 1254
QY 1192 SHSRSDNTFQDQKCHATTLMGSGFAQYQKGLQFQVNVGTGISAKMAEBSQKLRKA 1251
DB 1255 SHSRSDNTFQDQKCHATTLMGSGFAQYQKGLQFQVNVGTGISAKMAEBSQKLRKA 1314
QY 1252 INGVNMSVQFRLQGLQFQVNVGTGISAKMAEBSQKLRKA 1311
DB 1315 LHTGILQKTRBAGCGGCGELPHTGATRYFQKADYKRWVNTATPGLAEKRYMAGIKADYS 1374
QY 1312 FTPTDNISVYKZYFFVNVYVSNANVOTVNLTVLQPPGRYQKQVGLKABTLHFQISAF 1371
DB 1375 FKPAQHSITPFLSYSTIDNASGVTRVNTAVLAQDGKTRSMQVNAELKGFITSLH 1434
QY 1372 ISKSQSQQLKQKQNVKLGTRM 1394
DB 1435 AAKAQQLPQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1457

```

RESULT 9
 Q9F324 PRELIMINARY; PRT; 1454 AA.
 AC Q9F324;
 DT 01-MAR-2001 (TRENBERL 16, Created)
 DT 01-MAR-2001 (TRENBERL 16, last sequence update)
 DT 01-OCT-2003 (TRENBERL 25, last annotation update)
 DE Adhesion penetration protein.
 GN App.
 OS *Neisseria meningitidis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=24181;
 RA Ala-Aldeen D.A.;
 RL Submitted (0CT-2000) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=24181;
 RA Abdel Hadi H., Woodbridge K.G., Robinson K., Ala Aldeen D.A.;
 RT "Identification and Characterization of App: an Immunogenic
 Autotransporter Protein of *Neisseria meningitidis*,"
 RL Submitted (0CT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A0296277; CAC14671.1; --
 DR MEROPS; S06.006; --
 DR GO; GO:008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR006315; Autotransporter.
 DR InterPro; IPR005546; Autotransporter.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR000710; Peptidase_S6.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02395; IGAL; 1.
 DR Pfam; PF03212; Pertactin; 1.

SQ SEQUENCE 507 AA; 56170 MW; 665E365FEFE273F1 CRC64;
 Query Match 34.7%; Score 2523; DB 2; Length 507;
 Best Local Similarity 92.2%; Pred. No. 2.8e-101;
 Matches 471; Conservative 18; Mismatches 18; Indels 4; Gaps 1;

QY 70 KAPMIDFSVSRNGVAALVENQYIVSAHNVGYDVDFAGANNPDQHRFTYKIVKRNRY 129
 DB 1 KAPMIDFSVSRNGVAALVENQYIVSAHNVGYDVDFAGANNPDQHRFTYKIVKRNRY 60
 QY 130 KCONLHPYEDDYHNPRLHKEVTEAPIDMTSMNGSTYSDRKTYPERVIRIGSGRPWRND 189
 DB 61 KCONLHPYEDDYHNPRLHKEVTEAPIDMTSMNGSTYSDRKTYPERVIRIGSGRPWRND 120
 QY 190 QDKGQVAGAHYHLLTANTNNOGAGNGISYLGSDVRKAGYGEPLPIASGSGSSPMFI 249
 DB 121 QDKGQVAGAHYHLLTANTNNOGAGNGISYLGSDVRKAGYGEPLPIASGSGSSPMFI 180
 QY 250 YDAEKQKWLINGILREBNPEPKENGFOIVKRSYFDEIFERDLTSLYTRAGNGVYITISG 309
 DB 181 YDAEKQKWLINGILREBNPEPKENGFOIVKRSYFDEIFERDLTSLYTRAGNGVYITISG 240
 QY 310 NDNGQGSITQSGIPSEIKITLANNSLPLKEXKVHNPRIYDGNITSPRLNGETLYTMD 369
 DB 241 NDNGQGSITQSGIPSEIKITLANNSLPLKEXKVHNPRIYDGNITSPRLNGETLYTMD 300
 QY 370 QKQSLIFASDINGAGGLYFEGNFTVSPNSNQTWQAGIHVSNSYTWKVGVEHDL 429
 DB 301 QKQSLIFASDINGAGGLYFEGNFTVSPNSNQTWQAGIHVSNSYTWKVGVEHDL 360
 QY 430 SKIGKGLHVOAKGNGKSGISVGDQVILBQADDOGNKQAFSEIGLVSGRGVOLNDK 489
 DB 361 SKIGKGLHVOAKGNGKSGISVGDQVILBQADDOGNKQAFSEIGLVSGRGVOLNDK 420
 QY 490 QPDTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMIVNHTTQAAVNTTIGNESTVLPN 549
 DB 421 QPDTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMIVNHTTQAAVNTTIGNESTVLPN 476
 QY 550 GNNINKLDYRKELIANNVNGFETDKKNGRL 580
 DB 477 NDILKQLTNRKDIAFNMGFETDKKNGRL 507

RESULT 11
 Q8KRG5 PRELIMINARY; PRT; 505 AA.
 ID Q8KRG5
 AC Q8KRG5;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hap (Fragment)
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HK274, and HK75.
 RA Kilian M., Poulsen K., Lomholt H.;
 RT "Evolution of the paralogous hap and iga genes in Haemophilus
 RT influenzae: evidence for a conserved hap pseudogene associated with
 RT microcolony formation in the recently diverged Haemophilus aegyptius
 RT and H. influenzae aegyptius";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517141; AAM74927.1;
 DR EMBL; AF517142; AAM74928.1;
 DR GO; GO:0004295; P:trypsin activity; IEA.
 DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR000710; Peptidase_S6.
 DR Pfam; PF02395; IGA1; 1.
 DR PRINTS; PRO0921; IGASERPTASE.

DR PROSITE; PS00135; TRYPSIN_SER; 1.
 FT NON TER 1
 FT NON TER 505
 SQ SEQUENCE 505 AA; 55622 MW; ABA661F4CD54494 CRC64;
 Query Match 29.3%; Score 2128; DB 2; Length 505;
 Best Local Similarity 77.5%; Pred. No. 3e-84;
 Matches 396; Conservative 49; Mismatches 60; Indels 6; Gaps 2;

QY 70 KAPMIDFSVSRNGVAALVENQYIVSAHNVGYDVDFAGANNPDQHRFTYKIVKRNRY 129
 DB 1 KAPMIDFSVSRNGVAALVENQYIVSAHNVGYDVDFAGANNPDQHRFTYKIVKRNRY 60
 QY 130 KCONLHPYEDDYHNPRLHKEVTEAPIDMTSMNGSTYSDRKTYPERVIRIGSGRPWRND 189
 DB 61 KCONLHPYEDDYHNPRLHKEVTEAPIDMTSMNGSTYSDRKTYPERVIRIGSGRPWRND 120
 QY 190 QDKGQVAGAHYHLLTANTNNOGAGNGISYLGSDVRKAGYGEPLPIASGSGSSPMFI 249
 DB 121 QDKGQVAGAHYHLLTANTNNOGAGNGISYLGSDVRKAGYGEPLPIASGSGSSPMFI 180
 QY 250 YDAEKQKWLINGILREBNPEPKENGFOIVKRSYFDEIFERDLTSLYTRAGNGVYITISG 309
 DB 181 YDAEKQKWLINGILREBNPEPKENGFOIVKRSYFDEIFERDLTSLYTRAGNGVYITISG 240
 QY 310 NDNGQGSITQSGIPSEIKITLANNSLPLKEXKVHNPRIYDGNITSPRLNGETLYTMD 369
 DB 241 NDNGQGSITQSGIPSEIKITLANNSLPLKEXKVHNPRIYDGNITSPRLNGETLYTMD 300
 QY 370 QKQSLIFASDINGAGGLYFEGNFTVSPNSNQTWQAGIHVSNSYTWKVGVEHDL 429
 DB 301 QKQSLIFASDINGAGGLYFEGNFTVSPNSNQTWQAGIHVSNSYTWKVGVEHDL 360
 QY 430 SKIGKGLHVOAKGNGKSGISVGDQVILBQADDOGNKQAFSEIGLVSGRGVOLNDK 489
 DB 361 SKIGKGLHVOAKGNGKSGISVGDQVILBQADDOGNKQAFSEIGLVSGRGVOLNDK 420
 QY 490 QPDTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMIVNHTTQAAVNTTIGNESTVLPN 549
 DB 421 QPDTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMIVNHTTQAAVNTTIGNESTVLPN 476
 QY 550 GNNINKLDYRKELIANNVNGFETDKKNGRL 580
 DB 475 NDILKQLTNRKDIAFNMGFETDKKNGRL 505

RESULT 12
 Q8KSC0 PRELIMINARY; PRT; 510 AA.
 ID Q8KSC0
 AC Q8KSC0;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hap (Fragment)
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HK61;
 RA Kilian M., Poulsen K., Lomholt H.;
 RT "Evolution of the paralogous hap and iga genes in Haemophilus
 RT influenzae: evidence for a conserved hap pseudogene associated with
 RT microcolony formation in the recently diverged Haemophilus aegyptius
 RT and H. influenzae aegyptius";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517151; AAM74935.1;
 DR GO; GO:0004295; P:trypsin activity; IEA.
 DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR000710; Peptidase_S6.

Pfam: PF02395; IGAL: 1.
 DR PRINTS; PR000921; IGASERPHASE.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 FT MON TER 1 1
 FT NON TER 510 510
 SQ SEQUENCE 510 AA; 56795 MW; C5746C0C18949B4D CRC64;

Query Match 28.7%; Score 2087.5; DB 2; Length 510;
 Best Local Similarity 76.2%; Pred. No. 1,7e-82;
 Matches 390; Conservative 51; Mismatches 68; Indels 3; Gaps 2;

QY 70 KAPMIDFVSVRNGVAALVENQYIVSAHVAGTVDVGAAGNPNDRFTYKXKNY 129
 DB 1 KAPMIDFVSVRNGVAALVDQYIVSAHVAGTVDVGAAGNPNDRFTYKXKNY 60
 QY 130 KCONLHPEDDYENPRLHKEVTEAPIDMTSNMGSTYSDRKTYBRVRIGSGRFRMD 189
 DB 61 NHDKAHYLDYDHNPRHKEFTYDAAPIDMTSHDCKYANKKPKRYVVGSGDYMD 120
 QY 129 QDKGQVAGAYHYLTAGTTHNORAGNGSYLGQVRAKAGYGLPIAGSKDGSMPET 249
 DB 121 QNMRVYLSGYNLTGNTVNOGRSGDSIVRAGDIRKVDYGLPIAGSKDGSMPET 180
 QY 250 YDAKOKMLINGILREGNPFBAKENGQVLRKSYFDEIFERDLTSLYTRAGNGVYITSG 309
 DB 181 YDAFTQKMLINGILRAGQDYGEFDPQARSKSLDRIIRKQVNGFLPKNGVYITISK 240
 QY 310 NDNGGGSITQKSGITSEIKITLAMSILPKKQKTHNRNRYGRTYISPLNNGETLYEMD 369
 DB 241 SDGGIGAVTSKIKRPFEE--IPANNKLIKEDKDTYNNRYNGPIYISPOLNNGKXIFYGD 238
 QY 370 QKQGSILFASDINGAGGLYFEGNFTVSPNSNOTQAGHVSNSITVMKYNVEHDL 429
 DB 299 EELGSITLTIDGAGGLYFEGDPFTVSPNKETWKAGHVSISVTVMKYNVEHDL 358
 QY 430 SKIKGTLHVAQKENGKGSISVGDQKYLEEQADQSKKQAFSEIGLVSGRTVQLNDK 489
 DB 359 SKIKGTLHVAKENGKGSISVGDQKYLEEQADQSKKQAFSEIGLVSGRTVQLNDK 418
 QY 490 QEDTDKYFEGFRGRLDNLNGSLTFKRIQNTDEGAMIVNHTTQAAVITTNESIVLP 548
 DB 419 QEDTDKYFEGFRGRLDNLNGSLTFKRIQNTDEGAMIVNHTTQAAVITTNESIVLP 478
 QY 549 NGNNINKLDYRKEIAYNGFEGTDKNGKNGRL 580
 DB 479 NGNNINKLDYRKEIAYNGFEGTDKNGKNGRL 510

RESULT 13

QKRSB9 PRELIMINARY; PRT; 526 AA.

ID QKRSB9 PRELIMINARY; PRT; 526 AA.
 AC QKRSB9;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hap (Fragment).
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCB1_TaxID=727;
 RN NCB1_TaxID=727;
 RP SEQUENCE FROM N.A.
 RC STRAIN=HK715;
 RA "Kilian M., Poulsen K., Lomholt H.;
 RT "Evolution of the paralogous hap and iga genes in Haemophilus
 RT influenzae: evidence for a conserved hap pseudogene associated with
 RT microcolony formation in the recently diverged Haemophilus aegyptius
 RT and H. influenzae aegyptius.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517152; AAM74936.1; -;
 DR GO; GO:0004295; P:trypsin activity; IEA.
 DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Set_trypsin.

DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR000710; Peptidase_S6.
 DR Pfam; PF02395; IGAL: 1.
 DR PRINTS; PR000921; IGASERPHASE.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 FT MON TER 1 1
 FT NON TER 526 526
 SQ SEQUENCE 526 AA; 58159 MW; 4B77316D9E1B4F74 CRC64;

Query Match 25.3%; Score 1839.5; DB 2; Length 526;
 Best Local Similarity 66.7%; Pred. No. 9e-72;
 Matches 355; Conservative 52; Mismatches 98; Indels 27; Gaps 6;

QY 70 KAPMIDFVSVRNGVAALVENQYIVSAHVAGTVDVGAAGNPNDRFTYKXKNY 129
 DB 1 KAPMIDFVSVRNGVAALVDQYIVSAHVAGTVDVGAAGNPNDRFTYKXKNY 60
 QY 130 KCONLHPEDDYENPRLHKEVTEAPIDMTSNMGSTYSDRKTYBRVRIGSGRFRMD 189
 DB 61 KPGKQNPVHGYNHPRHKEFTYDAAPIDMTNDQKVANRDYPERVRIQSGHYWRD 120
 QY 190 QDK-----GQVAGAYHYLTAGTTHNORAGNGSYLGQVRAKAGYGLPIAGSKDGS 243
 DB 121 KQETNAVSSTDISGATNYLMAHTHTQSGDNGTTFSGATIRNNHYGRPIGAGQDS 180
 QY 244 GSPMFLYDAERQKMLINGILREGNPFBAKENGQVLRKSYF-----DEIFERDLHT 294
 DB 181 GSPMFLYDAERQKMLINGILREGNPFBAKENGQVLRKSYF-----DEIFERDLHT 236
 QY 295 SLYTRAGNGVYITTNNGGGSIT----QKSGITSEIKITLAMSILPKKQ--KYNPRY 349
 DB 237 --YIPISINSHSFPNSNDGKLTTRPSKQSKASVGTALKLPSELETAERAKAA 294
 QY 350 DGNPISPLNNGETLYEYDQKQGSILFASDINGAGGLYFEGNFTVSPNSNOTQAGI 409
 DB 295 PGNITQPMRHKXKITFGDRGTGLTENTINOGAGGLYFEGNFTVSPNSNOTQAGI 354
 QY 410 HVSNSITVMKYNVEHDLRLSKIKGTLHVAQKENGKGSISVGDQKYLEEQADQSKKQ 469
 DB 355 HVSNSITVMKYNVEHDLRLSKIKGTLHVAQKENGKGSISVGDQKYLEEQADQSKKQ 414
 QY 470 AFSEITLVSGSGRTVQLNDKQPTDXYFEGFRGRLDNLNGSLTFKRIQNTDEGAMIVN 529
 DB 415 AFSEITLVSGSGRTVQLNDKQPTDXYFEGFRGRLDNLNGSLTFKRIQNTDEGAMIVN 474
 QY 530 NTQAAVNTITGNESIVLP--NGNNINKLDYRKEIAYNGFEGTDKNGKNGRL 580
 DB 475 NTQAAVNTITGNESIVLP--NGNNINKLDYRKEIAYNGFEGTDKNGKNGRL 526

RESULT 14

QKRN0 PRELIMINARY; PRT; 514 AA.

ID QKRN0 PRELIMINARY; PRT; 514 AA.
 AC QKRN0;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hap (Fragment).
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCB1_TaxID=727;
 RN NCB1_TaxID=727;
 RP SEQUENCE FROM N.A.
 RC STRAIN=HK284; and HK292;
 RA "Kilian M., Poulsen K., Lomholt H.;
 RT "Evolution of the paralogous hap and iga genes in Haemophilus
 RT influenzae: evidence for a conserved hap pseudogene associated with
 RT microcolony formation in the recently diverged Haemophilus aegyptius
 RT and H. influenzae aegyptius.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517143; AAM74929.1; -;
 DR EMBL; AF517145; AAM74931.1; -;

DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR000710; Peptidase_S6.
 DR Pfam: PF02395; IGAL; 1.
 DR PRINTS: PR00921; IGASERPTASE.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 FT NON_TER 1 1
 FT NON_TER 514 514
 SQ SEQUENCE 514 AA; 56562 MW; 8848BDDF0A313 CRC64;
 Query Match 23.7%; Score 1724.5; DB 2; Length 514;
 Best Local Similarity 65.8%; Pred. No. 8e-67;
 Matches 342; Conservative 53; Mismatches 110; Indels 15; Gaps 8;
 QY 70 KAPMIDPSVSNAGVALVENQYIVSAHNVGYTDVDFGAEQNNPDQHFYKIVERNY 129
 DB 1 KAPMIDPSVSNAGVALVENQYIVSAHNVGYTDVDFGAEQNNPDQHFYKIVERNY 60
 QY 130 KQDNLHVEEDYHPRHLKFTYTAAPIMTGNMGSTYSPTKTYGERVIGSGRPFWRMD 189
 DB 61 KPGQ--XDGDTYHPRHLKFTYTAAPIMTGNMGSTYSPTKTYGERVIGSGRPFWRMD 117
 QY 190 --QDKGQ---VAGAYHYLTAGTNHQRGANGSYSLGQDYKXAGEYPLTAGSKGDS 243
 DB 118 EQAKADAVSSWLAGAIVNKLAKNTHVQGTGNGTVDLSGNLTKXNHGFLPIAGSGDS 177
 QY 244 GSDMFTYDAEKQKMLINGILIREGNPEKENGFPOLYKSYF-DEIFERDLTSLYTRAKN 302
 DB 178 GSDMFTYDAEKQKMLINGILIREGNPEKENGFPOLYKSYF-DEIFERDLTSLYTRAKN 237
 QY 303 GYTTISGNDNGQSSITQKSPETIKITLAMSILPKEDKYNRNYDGPVYSPLNNG 362
 DB 238 KHYLTAFANDNGTGYTL-KTEDSTSTTVKLEPITLSEGEVKEY--ARGANNFTKPKLDNG 294
 QY 363 ELYEYNDQKQSLIFASDINQAGALYFEKNFTVS-PNSNOTWQAGIHVSNSYTWKY 421
 DB 295 ELSIFIDQKQSLIFASDINQAGALYFEKNFTVS-PNSNOTWQAGIHVSNSYTWKY 354
 QY 422 NGVEHDLRSKIGKGLTHVQAKGKNGKSSISVDGKYLLEQADDOQKQASEICLYSGG 481
 DB 355 KNEPGRSLSKIGGLTHVQAKGKNGKSSISVDGKYLLEQADDOQKQASEICLYSGG 414
 QY 482 TYQANDKQPFQDFKYGFRGRLDNGSLTFKRIQNTQEGAMTVNHTTQAAVNTTG 541
 DB 415 TYQANDKQPFQDFKYGFRGRLDNGSLTFKRIQNTQEGAMTVNHTTQAAVNTTG 474
 QY 542 NESIVLP-NGANNINFLDYRKEIAYNGMFGEDTKXKNGRL 580
 DB 475 NESIVLP-NGANNINFLDYRKEIAYNGMFGEDTKXKNGRL 514
 RESULT 15
 Q9JXK3 PRELIMINARY; PRT; 1431 AA.
 AC Q9JXK3:
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Serine-type peptidase.
 GN NMB1998.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,

RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciccho A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Yamashiro Y.,
 RA Gill J., Scarlata V., Maignani V., Piza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.W., Moxon E.R., Rappelli R., Venter C.C.,
 RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58.";
 RL Science 247:1809-1815 (2000).
 DR EMBL: AF002549; AAF42325.1; -.
 DR PIR: A81018; A81018.
 DR TIGR: AMB1998; -.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR006315; Autotransporter.
 DR InterPro: IPR005546; Autotransporter.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR000710; Peptidase_S6.
 DR InterPro: IPR04899; Peptidase_S6.
 DR Pfam: PF03197; Autotransporter; 1.
 DR Pfam: PF02395; IGAL; 1.
 DR PRINTS: PR00921; IGASERPTASE.
 DR SMART: SM00020; ITP-SPC; 1.
 DR TIGRPFAM: TIGR01414; autotrans_bar1; 1.
 DR Complete proteome.
 SQ SEQUENCE 1431 AA; 157645 MW; 6801F0E209C24EFC CRC64;
 Query Match 23.0%; Score 1676; DB 16; Length 1431;
 Best Local Similarity 29.7%; Pred. No. 3.8e-64;
 Matches 470; Conservative 229; Mismatches 524; Indels 362; Gaps 50;

QY 10 FTACISISGI--VSQAMAGHTYFGIDQYRDEAKNKFTVGAQNTKYNKQGLVGS 67
 DB 9 FCSVSTIGLFAISPAVSSIVRNDVDQYFRDEAKNKGFVGAQNSISIDQKQGLVGS 68
 QY 68 MKKAPMDIDFSVSR-NGVALVENQYIVSAHNVGYTDVDFGAEQNNPDQHFYKIVERNY 126
 DB 69 LKGIEMPDPRVSRROFATLTVHPQYVNSVKNVGVGSJQGNLQMPBEQATYELVSR 128
 QY 127 NNYKKONLHPED-DYHNPRLKFTYTAAPIMTSS-----NNMGSTYCDRTKYERYR 178
 DB 129 N-----PPIYDHPRLKFTYTAAPIMTSS-----NNMGSTYCDRTKYERYR 180
 QY 179 ISSGRQFWRMDQKQDQVAGAYHYLTAGTNHQRGANGSYSLGQDYKXAGEYPLTAG 238
 DB 181 LSGTQYRKRADGRTTRTAPAYQVLTGTPPLKVLGPHNGHLLVGSLTQ-----PLNTYA 236
 QY 239 SKGDSGSMFTYDAEKQKMLINGILIREGNPEKENGFPOLYKSYFDEIFERDLTSL 296
 DB 237 IAGDGSPLFAFDHNRNFWLAGVL---STYAGFDFPKYIYVQPEFIRSTRQYETRL 293
 QY 297 YTRAGNGYVITSGNDNGQSSITQKSPETIKITLAMSILPKEDKYNRNYDGPVYSPLNNG 356
 DB 294 --DVALITNELIRKDNNGNSLTQ-GLAKERTLPIANPSL-APQNDRAHMPSDA----- 344
 QY 357 PRLANGERTLVE--MDQKQSLIFASDINQAGALYFEKNFTVS-PNSNOTWQAGIHVS 413
 DB 345 -----GKTLISSREPK--TLMADNINQAGALQFDSNFTV-VGKHTWQAGIYVAD 396
 QY 414 NSTVTWKQNGVEHDLRSKIGKGLTHVQAKGKNGKSSISVDGKYLLEQADDOQKQASE 473
 DB 397 GKRVFQVQNSKGRSLKIGAGTLIANQGINQDSSISDEGTVVLAKQAASDQSKQAFQ 456
 QY 474 IGVSARGVQVQNSKGRSLKIGAGTLIANQGINQDSSISDEGTVVLAKQAASDQSKQAFQ 533
 DB 457 VGLTSRGVQVQNSKGRSLKIGAGTLIANQGINQDSSISDEGTVVLAKQAASDQSKQAFQ 516
 QY 534 AANVTITANE-----SIYPNKR----- 551
 DB 517 AATLITITGNPILSPEHVEWQNGNRPQNAAYEYINPRNRRTDYFLKPGNPREPEP 576
 QY 552 -----NINLDRKRLA-----YNGMFGEDTKXKNGRL 571

Db 577 LMKONSTWOFIGNRQQAEOAENAPBDLITPGYIGENAOGTGAAPSYKTNMEA 636
 QY 572 -DNKX-NGRLNLKXPTEDRTLLSGTNLKGDIOTKGLPPSG 616
 Db 637 IEKTRHANAVYKPEYRNGALNLYRPRKRDSTLLNGMNLNGEVLIGCMWISG 696
 QY 617 RPTTHAYNLNKRNSMEGIPQGETVMDHMINRTFKENFOIKGSAYVS-RVYSSLEG 675
 Db 697 RPYHAJDHQA-R-EPVLENEWTDGSSFKARFTLNARLTAGRNPAHDG 746
 QY 676 NMTYSNANATPGVNPNOONTICTRSDWGLTTCQKVDLTDTKYNSIPKQINGSLT 735
 Db 747 DITAYDLSGIDJFTOG-KTPECTRYSHSGTHCTPNAYLKAENYALPATQVRGDITIN 805
 QY 736 DNATANYKGLAKINGVTLTNHSGFTLSNNAOIGNIRLSDNSTATVDNANINENHJTD 795
 Db 806 DRS-ELRL-GKALYG 819
 QY 796 SNOFSLNKHSHQIQDKGTYTLE-NATWTPSDTLQNLTLNNGSTLLNSAYSASN 854
 Db 820 S-IRAGDPAVREADSNWTLISGSHTGALTLIDGAQITLNPDPANNTH 866
 QY 855 NTPRRSLETTETPTSABRPNLTVNGKLSGGTFOF-TSSLFGYKSDKLKLSNDAEGD 913
 Db 867 N-NRFTLTVNGTLDGEGTFRFLGTIVAKQAPPLKLGDSRGA 909
 QY 914 YLISVRNTEKEPETLEBOLTIVESKNOPLSDKLFTLENDVADGALRYLKNDEFR 973
 Db 910 FQIHVKWGOEPQTESLAVSL-NEKSHQARFTLQNGYADIGAYRYLKRNNNGYSL 967
 QY 974 HNPKEQELNDLVRAE-QAERTLEAKOVEPTAKTGTGEPKVSRAARAAPDTL 1028
 Db 968 YNPLEKALQTEATRABERNOQANQOATDI-SROYHDSD 1009
 QY 1029 POSILNLEAKOAEIT-AETQSNKATKYKRAVFSDDPLD-QSL-F 1375
 Db 1010 ATRQALQAMQNSQTEHARIDSQVYLSAKQ-TDPLTGILTRAQNLCAQGY 1061
 QY 1076 ALEALAEVTDAPQOS-EKDRLEAKOEAKOROKD- 1108
 Db 1062 SADICQVAKAATNDLTFETELDTYIEVENABEELDKARQGGDAQAVETARHAYLNA 1121
 QY 1109 -LISRYNSALSELGATVNSMLSVODELDRLEFVDA 1143
 Db 1122 LMRLSRQJHSIKTGVAIGIRMPYLAHLISRSNATVSEQAAYRTGQOAGRIDRHLDPQ 1181
 QY 1144 QSAVWNTLAQDKERYSDAFRAYOQOKTNLRQIGVQKALANG-RIGAVFSHSSDNTPE 1202
 Db 1182 QONWLEGTQNDYHSGTHRY-QQTTNVAHIGIQTGIDRLSVGITLDERTNRPDE 1240
 QY 1203 QV-KNAKILTMNSGPAQYQWMDLOFGVNTGTGISAKAEQSRKIHRAJNYGVN 1257
 Db 1241 GVSANRNSGAHL-FVKGENALFMAADGYNSRTRFTYDGAAVRRHAMDAGIN 1295
 QY 1258 ASYQFLGQIGIOPYGVNRYFIEEN-YOSEEVAVTPSIAFNRYNAGIRVDYTFP 1314
 Db 1296 TGKIDTGS-INTRPYAGIR-INSNGKRYVLDGALINSPAQIOTTMHGIRLDKT-VE 1350
 QY 1315 TNSISVKEPYFPVNYDV-SNANVQTTVN-LTVLQOPFG-RYWKQEVGLKAEILHFOIS 1369
 Db 1351 LQOAKLTPAFSSDYVHPRONGSALSVNDRILLQOAHGTLHTLQIDAGYK-WNAK 1406
 QY 1370 APTSKSGSGLGRKQONVGYLGRW 1394
 Db 1407 LHAAYGDSNTARHKOAGIKGINW 1431

1	7273	100.0	1394	2	AAR2768	Aar92768	Adnestrion
2	7273	100.0	1394	6	ADA09338	Ada09343	Haemophil
3	7246.5	99.6	1395	6	ADA09338	Ada09348	Haemophil
4	61.5	84.1	1434	6	ADA09350	Ada09350	Haemophil
5	5827.5	80.1	1412	6	ADA09358	Ada09358	Haemophil
6	5807	79.8	1436	6	ADA09354	Ada09354	Haemophil
7	5303.5	72.9	1391	6	ADA09356	Ada09356	Haemophil
8	5297.5	72.8	1391	6	ADA09352	Ada09352	Haemophil
9	4178	57.4	1449	6	ABU37787	Abu37787	Protein e
10	4090.5	56.2	1454	3	AA56661	Aa56621	Netseria
11	4085	56.2	1457	3	AA58852	Aa58852	N. mening
12	4085	56.2	1457	4	AAU27556	Aa27556	Netseria
13	4085	56.2	1457	6	ABU08601	Abu08601	Netseria
14	4083	56.1	1457	3	AA56622	Aa56622	Netseria
15	4082	56.1	1457	3	AA56562	Aa56562	N. mening
16	4076	56.0	1457	2	AA58823	Aa58823	Netseria
17	4072.5	56.0	1468	6	ABP77229	Abp77229	N. gonorr
18	4072.5	56.0	1468	6	ABU37103	Abu37103	Protein e
19	4069	55.9	1449	2	AA58824	Aa58824	Netseria
20	4061.5	55.8	1448	2	AA58825	Aa58825	Netseria
21	4044.5	55.6	1439	4	AAU27557	Aa27557	Netseria
22	3409.5	46.9	1178	6	ABU08603	Abu08603	Netseria
23	3354.5	46.1	1136	6	ABU08605	Abu08605	Netseria
24	2981	40.4	956	6	ABU08602	Abu08602	Netseria
25	2846	39.7	914	6	ABU08604	Abu08604	Netseria

UNITED STATES DEPARTMENT OF JUSTICE
FEDERAL BUREAU OF INVESTIGATION
WASHINGTON, D.C. 20535

PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX St Geme JW, Falkow S;
 XX WPI, 1996-151147/15.
 DR N-PSDB; AAT17215.
 XX Haemophilus adhesion and penetration protein and corresponding DNA - used
 PT to produce vaccines against H. influenzae infection.
 XX Claim 2, Fig 6, 105pp; English.
 XX The sequence represents a Haemophilus influenzae adhesion and penetration
 CC protein, encoded by the hap gene. The protein is first synthesized as a
 CC preprotein with 3 functional domains (the N-terminal signal peptide, the
 CC protease and a C-terminal helper domain), which is transported to the
 CC periplasm, followed by insertion of the C-terminal beta-domain into the
 CC outer membrane, possibly forming a pore, and export of the N-terminal
 CC fragment through the outer membrane, followed by autolytic cleavage
 CC and secretion of the mature protease, leaving an outer membrane protein
 CC fragment. An active site serine residue characteristic of serine
 CC proteases is present, along with 2 cysteine residues which may be
 CC important in forming the catalytic domain. A gene encoding the protein
 CC may be inserted in a vector and expressed in recombinant host cells, for
 CC use as a recombinant vaccine. The protein may also be used in diagnostic
 CC monoclonal antibody production.
 XX Sequence 1394 AA;
 SC
 Query Match 100.0%; Score 7273; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 601 LKGDITQTKGLPFSSGRPTPHAYNHLNKRMSMEGIPOGELIWDHDMINRTFAENFOIK 660
 QY 661 GGSAAVSSNVSIEGWTVTSNAAATGCVVPNOQNTICTRSMTGLTTCOKVDLDTKYI 720
 DB 661 GGSAAVSSNVSIEGWTVTSNAAATGCVVPNOQNTICTRSMTGLTTCOKVDLDTKYI 720
 QY 721 NSIPKQINSINLTDAATANKYGLAKINGVTLTHSQFTLSNNAQIGNILSDNSTA 780
 DB 721 NSIPKQINSINLTDAATANKYGLAKINGVTLTHSQFTLSNNAQIGNILSDNSTA 780
 QY 781 TVDNANLNGVHLTDSAPSLANSHFSHOIQDGKTTVLTENATWTPSDTLQNTLNN 840
 DB 781 TVDNANLNGVHLTDSAPSLANSHFSHOIQDGKTTVLTENATWTPSDTLQNTLNN 840
 QY 841 STTLNAYSASSNTPRRRSLETTTTSAREPNTLVNKGSLGGCTQFTSSIFGK 900
 DB 841 STTLNAYSASSNTPRRRSLETTTTSAREPNTLVNKGSLGGCTQFTSSIFGK 900
 QY 901 SDKLKLSNDAGDYILSVNTGKEPTELQGLTVESKDNQPLSDKLKPTLENHVDAGL 960
 DB 901 SDKLKLSNDAGDYILSVNTGKEPTELQGLTVESKDNQPLSDKLKPTLENHVDAGL 960
 QY 961 RYKLVKNDGEFRLLNPIKQELNDLVRAEQARTLEAKQVETPAKQTEBPYKRRRA 1020
 DB 961 RYKLVKNDGEFRLLNPIKQELNDLVRAEQARTLEAKQVETPAKQTEBPYKRRRA 1020
 QY 1021 RAAPETPLPDSSLNLEAKQELTAETQSKAKTKKVSRAVSPDLDQSLPTEEA 1080
 DB 1021 RAAPETPLPDSSLNLEAKQELTAETQSKAKTKKVSRAVSPDLDQSLPTEEA 1080
 QY 1081 LEVIDAPOOSEKDLAQEAEKQKQKOLISRSNSALSELATVNSVLGVDELRLTV 1140
 DB 1081 LEVIDAPOOSEKDLAQEAEKQKQKOLISRSNSALSELATVNSVLGVDELRLTV 1140
 QY 1141 DQASAVNTNIAQDKRRYDSDA FRAYQOQKTNLROIGVAKALANGRI GAVFSRSNDTE 1200
 DB 1141 DQASAVNTNIAQDKRRYDSDA FRAYQOQKTNLROIGVAKALANGRI GAVFSRSNDTE 1200
 QY 1201 DEGVNNAHTLTMGSGPQYQWGDLOFGNVGTGTSASVNAEESQKTHRAIVNGNASY 1260
 DB 1201 DEGVNNAHTLTMGSGPQYQWGDLOFGNVGTGTSASVNAEESQKTHRAIVNGNASY 1260
 QY 1261 QFTLQGLGIQYFGVNRYPFERENYQSEHVAVKTPSLAENRYNAGISVDYFTPTNISV 1320
 DB 1261 QFTLQGLGIQYFGVNRYPFERENYQSEHVAVKTPSLAENRYNAGISVDYFTPTNISV 1320
 QY 1321 KPYFPVNVVDVSNNAVQTTNLTVLQOPFRYQKEVGLKAEIILHFOISAFISKQSQSL 1380
 DB 1321 KPYFPVNVVDVSNNAVQTTNLTVLQOPFRYQKEVGLKAEIILHFOISAFISKQSQSL 1380
 QY 1381 GKOQNVGYKLGEM 1394
 DB 1381 GKOQNVGYKLGEM 1394

RESULT 2
 ADA09343
 ID ADA09343 standard; protein; 1394 AA.
 XX ADA09343;
 AC 06-NOV-2003 (first entry)
 DT XX
 DE Haemophilus adherence and penetration protein (HAP).
 XX Haemophilus adhesion and penetration protein; HAP; antibacterial;
 KW immunostimulant; vaccine; Haemophilus influenza infection; influenza.
 OS Haemophilus influenzae.
 XX US2003073166-A1.
 XX

PD 17-APR-2003.
 XX
 PF 22-FEB-2002; 2002US-00080505.
 XX
 PR 25-AUG-1994; 94US-00296791.
 PR 20-APR-2001; 2001US-0083996.
 XX
 PA (GEMV/) GEMV J W S.
 XX
 PI Gene JMS;
 XX
 XX WPI; 2003-567308/53.
 DR N-PSDB; ADA09342.
 XX
 PT New recombinant Haemophilus adhesion and penetration protein, useful as a
 PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
 PT influenzae infections, in generating antibodies for passive immunization.
 PS
 PS Example 1, Fig 6; 149p; English.
 CC The invention relates to a new recombinant Haemophilus adhesion and
 CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
 CC ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
 CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
 CC acid that will hybridize under stringency conditions to the nucleic acid
 CC detailed above, an expression vector comprising transcriptional and
 CC translation regulatory nucleic acid operably linked to the HAP nucleic
 CC acid, producing HAP by culturing a host cell transformed with the
 CC expression vector, a composition comprising a pharmaceutical carrier and
 CC HAP, inducing an immune response in a patient by administering the
 CC composition, a composition comprising an antibody capable of binding to
 CC HAP (or to a peptide selected from 35 peptide sequences appearing as
 CC ADA09359-ADA09393), and an antibody capable of binding to the peptides.
 CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
 CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
 CC infection (e.g. influenza), in generating an immune response in a
 CC patient, and in generating antibodies to the HAP protein. The antibodies
 CC may be used to diagnose the presence of H. influenzae infection in a
 CC sample or patient, in passive immunisation treatments, and to purify or
 CC separate HAP proteins or the H. influenzae organism from a sample. The
 CC present sequence is the H. influenzae HAP protein.
 XX
 XX Sequence 1394 AA:
 SQ
 Query Match 100.0%; Score 7273; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 361 NGETLYFMDQKQSLIFASDINQAGGLYFEGNFVSPRSNQTWQAGIHVSENSTVTK 420
 QY 421 VNGVEHDELSKIGKGLTHVQAKGENKKSISVGGKVIIEQQADQGNKQAFSEIGVSGR 480
 DB 421 VNGVEHDELSKIGKGLTHVQAKGENKKSISVGGKVIIEQQADQGNKQAFSEIGVSGR 480
 QY 481 GTVOLNDKQFDYDKFYFGRGRLDNGSLTPKRIQNTDEGMYVNNNTQANVTIT 543
 DB 481 GTVOLNDKQFDYDKFYFGRGRLDNGSLTPKRIQNTDEGMYVNNNTQANVTIT 540
 QY 541 GNESTVLPNGNINIKLDYRKEIAYNGWPGETDKKNGKLNLIYKPTTEORTLLSGGN 600
 DB 541 GNESTVLPNGNINIKLDYRKEIAYNGWPGETDKKNGKLNLIYKPTTEORTLLSGGN 600
 QY 601 LKGDITQTKGLFSGRPPTPHAYNHLNKRSEMEGIPOGEIYVMDHMINTPFAENFOIK 660
 DB 601 LKGDITQTKGLFSGRPPTPHAYNHLNKRSEMEGIPOGEIYVMDHMINTPFAENFOIK 660
 QY 661 GSAVAVSRVSSIBGWWTVSNNAATPGVVPNQONTICTRSDMTGLTTCOKVDLTKYI 720
 DB 661 GSAVAVSRVSSIBGWWTVSNNAATPGVVPNQONTICTRSDMTGLTTCOKVDLTKYI 720
 QY 721 NSIPRTQNGSINTDNATANKGIAKINGVTLTNHSQFTLSNNAQOIGIRISDNSTA 780
 DB 721 NSIPRTQNGSINTDNATANKGIAKINGVTLTNHSQFTLSNNAQOIGIRISDNSTA 780
 QY 781 TVDNANLNGVHLTDSNOFSLNHSRSHQIQGKTTVTLBNATVMSDITLQNTLNN 840
 DB 781 TVDNANLNGVHLTDSNOFSLNHSRSHQIQGKTTVTLBNATVMSDITLQNTLNN 840
 QY 841 STTLNLSAVSASNTPTRRRSLETTTTSAREPNTLVNKLSSGQTFQTFSSILFGY 900
 DB 841 STTLNLSAVSASNTPTRRRSLETTTTSAREPNTLVNKLSSGQTFQTFSSILFGY 900
 QY 901 SDKLKSNDAAGDYTLISYNTGKEPETEQLTVESQNDQPSDKLKTLENDHVAGAL 960
 DB 901 SDKLKSNDAAGDYTLISYNTGKEPETEQLTVESQNDQPSDKLKTLENDHVAGAL 960
 QY 961 RYVLVNDGEERLNPRIKBOELHNDLVABQERTLEANKVPEPTAKTGTGEPKYSRPA 1020
 DB 961 RYVLVNDGEERLNPRIKBOELHNDLVABQERTLEANKVPEPTAKTGTGEPKYSRPA 1020
 QY 1021 RAAPPTLPDQSLNALNAQOELTPEOKSKATTKVSKRAVSDPLDQSLFALEA 1080
 DB 1021 RAAPPTLPDQSLNALNAQOELTPEOKSKATTKVSKRAVSDPLDQSLFALEA 1080
 QY 1081 LEVINDAPQOESKRLAQEBAEKQKQKDLISRYNSALSLSATVNSMLSYQDELRLFY 1140
 DB 1081 LEVINDAPQOESKRLAQEBAEKQKQKDLISRYNSALSLSATVNSMLSYQDELRLFY 1140
 QY 1141 DQASAVWNTLADOKRRYDSDAFRAYQOQKTNLRQIGVQKALANGRIGAVFSHSRDNTE 1200
 DB 1141 DQASAVWNTLADOKRRYDSDAFRAYQOQKTNLRQIGVQKALANGRIGAVFSHSRDNTE 1200
 QY 1201 DBOYKKAHLTMGSGPYQYQMGDLOPGVNVGTGISASKAAEFSQSKIRKAINGVNASY 1260
 DB 1201 DBOYKKAHLTMGSGPYQYQMGDLOPGVNVGTGISASKAAEFSQSKIRKAINGVNASY 1260
 QY 1261 QFRIGQIGIOPYGVNRYFERENYQSEBEVYKPSLAENRYAGIRVDTYFTPTDNISY 1320
 DB 1261 QFRIGQIGIOPYGVNRYFERENYQSEBEVYKPSLAENRYAGIRVDTYFTPTDNISY 1320
 QY 1321 KPYEFYNYDVSNNAVQTVNLTLVQPPGRYQWQEVGLKAEILHFOISAFISYSGSSQL 1380
 DB 1321 KPYEFYNYDVSNNAVQTVNLTLVQPPGRYQWQEVGLKAEILHFOISAFISYSGSSQL 1380
 QY 1381 GKQONNGVYLGTRY 1394
 DB 1381 GKQONNGVYLGTRY 1394

RESULT 3

Office

Sequence 1395 AA;

Query Match	99.6%	Score 7246.5	DB 6	Length 1395
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1393; Conservative	0	Mismatches	0	Indels 3; Gaps 2

Qy	1	MKKVGRNLPLFLACISLGIYSQAMAGHYIFGIDVQYRPBARKKGFVGAQNIKYTKNO	60
Qy	1	MKKVGRNLPLFLACISLGIYSQAMAGHYIFGIDVQYRPBARKKGFVGAQNIKYTKNO	60
Db	1	MKKVGRNLPLFLACISLGIYSQAMAGHYIFGIDVQYRPBARKKGFVGAQNIKYTKNO	60
Qy	61	GOLVGRSMTKAMLIDFSGVSRNGVALVENQYIYSVAHNGVTDVDFGAEGNPPQJHRT	120
Qy	61	GOLVGRSMTKAMLIDFSGVSRNGVALVENQYIYSVAHNGVTDVDFGAEGNPPQJHRT	120
Db	61	GOLVGRSMTKAMLIDFSGVSRNGVALVENQYIYSVAHNGVTDVDFGAEGNPPQJHRT	120
Qy	121	YKIVGRNNYKQNLHFEDDYNRPRLHKEVTEAAPIDMTSNNGSTYSIDRTKPEERVRIG	180

Dh	121	YK1VKNRYKKNDLHFEYDDYHNPDLKFEVTEAPADIMTSMNMGSTYSDEKTKPERVRIG	180
Qy	181	SGQCFRRDDKCDQJAGAHYHLLTAKTNNQAGANCYSTRJGDYRKAAGYGLPLAAGSK	240
Dh	181	SGCFRRDDQJCDQJAGAHYHLLTAKTNNQAGANCYSTRJGDYRKAAGYGLPLAAGSK	240
Qy	241	GDSGPMFIYDAEKQWLLNGILBEGNPEEGKXKPOLYKSPFELFEEDLHLSLYTRA	300
Dh	241	GDSGPMFIYDAEKQWLLNGILBEGNPEEGKXKPOLYKSPFELFEEDLHLSLYTRA	300
Qy	361	NGEYTYFENDQKQSLIPASDINOGAGLFEBCNPLVSPDNNQWQAGIHVSENSVTWK	420
Dh	361	NGEYTYFENDQKQSLIPASDINOGAGLFEBCNPLVSPDNNQWQAGIHVSENSVTWK	420
Qy	421	VNVEZHDLSKIKGKGLTHVQKGNKKSISVGDQVLIIEQADQGNKQAFSEIGSVGR	480
Dh	421	VNVEZHDLSKIKGKGLTHVQKGNKKSISVGDQVLIIEQADQGNKQAFSEIGSVGR	480
Qy	481	GYVQJLNDKQJDDPKYFEPFGRGLNHSLLTPKQJOMDEBAMIVNKHQTQAAWVIT	540
Dh	481	GYVQJLNDKQJDDPKYFEPFGRGLNHSLLTPKQJOMDEBAMIVNKHQTQAAWVIT	540
Qy	541	GNESYVLBNGNINIKLDYKELIAYNGFEDTJKNNGRNLNLYPTEDEPILLSGGTN	600
Dh	541	GNESYVLBNGNINIKLDYKELIAYNGFEDTJKNNGRNLNLYPTEDEPILLSGGTN	600
Qy	601	LKQDITQKGLFESSRPLPHAYVHLMKKSMBGIPQGEIYVDBHDINRPFKAEPOLK	660
Dh	601	LKQDITQKGLFESSRPLPHAYVHLMKKSMBGIPQGEIYVDBHDINRPFKAEPOLK	660
Qy	661	GSGAVSBNVSIIEGMWTSNNANATGVVPQWITCTESBWTGLTQCKVLDITTKYI	720
Dh	661	GSGAVSBNVSIIEGMWTSNNANATGVVPQWITCTESBWTGLTQCKVLDITTKYI	720
Qy	721	NSIPKTXNGSINLTQNACTAVYKGLAKNGVTLTNSQFTLSNNAIQINRLSNSTA	780
Dh	721	NSIPKTXNGSINLTQNACTAVYKGLAKNGVTLTNSQFTLSNNAIQINRLSNSTA	780
Qy	781	TJUNNANLGNVHLTDSAQSLKNSHSHQIQDGGTIVLENAVWMPSDTLQNTLNN	840
Dh	781	TJUNNANLGNVHLTDSAQSLKNSHSHQIQDGGTIVLENAVWMPSDTLQNTLNN	840
Qy	841	STTTNSAYSASSMNP--RRSLETPETPSAEHRFNTLTVNGKLSQGFQFTSILFG	898
Dh	841	STTTNSAYSASSMNP--RRSLETPETPSAEHRFNTLTVNGKLSQGFQFTSILFG	898
Qy	899	YYSKDKLISNDASDYLILSVNTGEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	958
Dh	899	YYSKDKLISNDASDYLILSVNTGEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	958
Qy	959	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	960
Dh	959	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	960
Qy	999	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1000
Dh	999	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1000
Qy	1001	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1002
Dh	1001	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1002
Qy	1003	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1004
Dh	1003	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1004
Qy	1005	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1006
Dh	1005	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1006
Qy	1007	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1008
Dh	1007	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1008
Qy	1009	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1010
Dh	1009	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1010
Qy	1011	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1012
Dh	1011	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1012
Qy	1013	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1014
Dh	1013	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1014
Qy	1015	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1016
Dh	1015	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1016
Qy	1017	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1018
Dh	1017	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1018
Qy	1019	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1020
Dh	1019	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1020
Qy	1021	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1022
Dh	1021	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1022
Qy	1023	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1024
Dh	1023	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1024
Qy	1025	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1026
Dh	1025	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1026
Qy	1027	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1028
Dh	1027	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1028
Qy	1029	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1030
Dh	1029	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1030
Qy	1031	ALSYKYLKNSDAZEDYILSVNRTGKEPTELEOLIVYSSKNOPLSDKKFTLKNKPYDAG	1032
Dh	10		

QY 1259 SYGFRLLGGLQIPYFVGVRVIFERENYQSEERVKTSLAPNRNAGIRVDYETPTDNI 1318
 DB 1260 SYGFRLLGGLQIPYFVGVRVIFERENYQSEERVKTSLAPNRNAGIRVDYETPTDNI 1319
 QY 1319 SYVPYFVNVVDSNANVQTTNLTVLQOPFGRRYMOKEVGLAKELHFOISAFISQSS 1378
 DB 1320 SYVPYFVNVVDSNANVQTTNLTVLQOPFGRRYMOKEVGLAKELHFOISAFISQSS 1379
 QY 1379 QLGKQONVGVKLGTRM 1394
 DB 1380 QLGKQONVGVKLGTRM 1395
 RESULT 4
 ADA09350
 ID ADA09350 standard; protein; 1434 AA.
 XX ADA09350;
 AC
 XX
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Haemophilus HAP protein from strain 11.
 XX
 XX Haemophilus adherence and penetration protein; HAP; antibacterial;
 KM immunostimulant; vaccine; Haemophilus influenza infection; influenza.
 XX
 OS Haemophilus influenzae; strain 11.
 PN US2003073166-A-
 XX
 PD 17-APR-2003.
 XX
 XX 22-FEB-2002; 2002US-00080505.
 PR 25-AUG-1994; 94US-00296791.
 PR 20-APR-2001; 2001US-0083996.
 XX
 XX (GENE/) GENE J W S.
 PA
 XX
 P1 Gene JMS;
 DR WPI; 2003-567308/53.
 DR N-PSDB; ADA09349.
 XX
 PT New recombinant Haemophilus adhesion and penetration protein, useful as a
 PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
 PT influenzae infections, in generating antibodies for passive immunization.
 XX
 PS Claim 2; Fig 17; 149pp; English.
 XX
 CC The invention relates to a new recombinant Haemophilus adhesion and
 CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
 CC ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
 CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
 CC acid that will hybridize under stringency conditions to the nucleic acid
 CC detailed above, an expression vector comprising transcriptional and
 CC translation regulatory nucleic acid operably linked to the HAP nucleic
 CC acid, producing HAP by culturing a host cell transformed with the
 CC expression vector, a composition comprising a pharmaceutical carrier and
 CC HAP, inducing an immune response in a patient by administering the
 CC composition, a composition comprising an antibody capable of binding to
 CC HAP (or to a peptide selected from 35 peptide sequences appearing as
 CC ADA09359-ADA09363), and an antibody capable of binding to the peptides.
 CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
 CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
 CC infection (e.g. influenza), in generating an immune response in a
 CC patient, and in generating antibodies to the HAP protein. The antibodies
 CC may be used to diagnose the presence of H. influenzae infection in a
 CC sample or patient, in passive immunisation treatments, and to purify or
 CC separate HAP proteins or the H. influenzae organism from a sample. The
 CC present sequence is the HAP protein from a strain of H. influenzae.

SQ Sequence 1434 AA;
 Query Match 84.1%; Score 6115; DB 6; Length 1434;
 Best local similarity 82.9%; Pred. No. 0;
 Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 10;
 QY 1 MKCTVFRILNPLTACISLGIVSQAAGHTTYGIDYOTRDPAENKGTTCGAQIKYVTKQ 60
 DB 1 MKCTVFRILNPLTACISLGIVSQAAGHTTYGIDYOTRDPAENKGTTCGAQIKYVTKQ 60
 QY 61 GOLVGTSMTRAPMIDSVSVSRNGVAALVENQYIVSAHVNGVTVVDGARGNPDQHRFT 120
 DB 61 GOLVGTSMTRAPMIDSVSVSRNGVAALVENQYIVSAHVNGVTVVDGARGNPDQHRFT 120
 QY 121 YKIVKNNYKKDMIAHYEDDYNNPRLHKFTTEAPITDMTSMNKGSTYSDRTKYPERVIG 180
 DB 121 YKIVKNNYNNHDAKHRYLDDYNNPRLHKFTTDAPIDMTSMNKGSTYSDRTKYPERVIG 180
 QY 181 SGRQFWRNDQDKGDVAGAYHYTLNNTNQGAGNGISYLAGDVRKAGETGPLTASSK 240
 DB 181 SGRQFWRNDQDKGDVAGAYHYTLNNTNQGAGNGISYLAGDVRKAGETGPLTASSK 240
 QY 241 GDSGPMFTIDAEQMLNGLVLRBQPYGDFQGLAKKSFIDEIRDOQNGFLTPK 300
 DB 241 GDSGPMFTIDAEQMLNGLVLRBQPYGDFQGLAKKSFIDEIRDOQNGFLTPK 300
 QY 301 GNGVYTTISGNDGQSGISITQKSGISEIKITLANMSLPLKEXDKVHNPYDGPNTYSPRLN 360
 DB 301 GNGVYTTISKDDGIGVYVTSKIGKPRE--IPLANNKLIKEDQYVNNRNGPNTYSPQLN 358
 QY 361 NSETLYFMDQKQSLIPASDINQAGGLYFENFTYSPNSQNTMQAGCHVSNSTVYWK 420
 DB 359 NKNITFQDEELASITLITLTDIOGAGGLYFEGDFIVSPKMEKWKAGIHVSEISTVYWK 418
 QY 421 VNGVEHPLSLKIGKGLVQAKNGKSGISVSDKYLTEQADQKQKQAFSEITGVYSR 480
 DB 419 VNGVENDRLSKIGKGLVQAKNGKSGISVSDKYLTEQADQKQKQAFSEITGVYSR 478
 QY 481 GTVOLNDKQPDPTDKFYFGRGGRGLDNGHSITFKKIQNTDEGAMVYVNTHTQAAVNTT 540
 DB 479 GTVOLNDKQPDPTDKFYFGRGGRGLDNGHSITFKKIQNTDEGAMVYVNTHTQAAVNTT 538
 QY 541 GNESTVLP--NKNINIKLYRKEIAYNGFSGEDDNKNGRLNLKPTEDRLLLSGCT 599
 DB 539 GNESTVLP--NKNINIKLYRKEIAYNGFSGEDDNKNGRLNLKPTEDRLLLSGCT 598
 QY 600 NIKGPIITQKGLPFESGRPTPHAVYHLNKRSEMEGIPQGEIWMDDHNPRTKFAENPOI 659
 DB 599 NIKGPIITQKGLPFESGRPTPHAVYHLNKRSEMEGIPQGEIWMDDHNPRTKFAENPOI 658
 QY 719 INSISPTQINGSLINDNATANKGLAVKAGNTTLNHSQFTLSNATQIGNRIS-- 775
 DB 719 INSISPTQINGSLINDNATANKGLAVKAGNTTLNHSQFTLSNATQIGNRIS-- 778
 QY 776 ---DST-----ATVDNANGVNHITDSAOFS 800
 DB 775 ARVNNATLMDGVNLADTREFSLSNQATQIGTSLHQQAQVATVDNANGVNHITDSAOFS 838
 QY 801 LKNSHSHQIGDKGTTVTLENATWTPSDTTLQNTLNNSTITLNSAYSASSNTP--RR 859
 DB 839 LKNSHSHQIGDKGTTVTLENATWTPSDTTLQNTLNNSTITLNSAYSASSNTP--RR 858
 QY 860 KSLFETTPTAHERNTTLVNGKXSGGTQFTSLIFGKSCDKLSNDAEGDYLLTSVR 919
 DB 859 KSLFETTPTAHERNTTLVNGKXSGGTQFTSLIFGKSCDKLSNDAEGDYLLTSVR 958
 QY 920 NTKCEFTLEQLTVESKONPLSDKLFTELNDHDAQALRYKLVKDGSEFRLLNPDIK 979
 DB 959 NTKCEFTLEQLTVESKONPLSDKLFTELNDHDAQALRYKLVKDGSEFRLLNPDIK 1018

QY 781 TVDNAMNGNWHITDSAQFSLKNHSHQIQDGKTVTLLENNVTWMSDTLLQNTLNN 840
 DB 773 TVDNAMNGNWHITDSAQFSLKNHSHQIQDGKTVTLLENNVTWMSDTLLQNTLNN 832
 QY 841 STTINSAYASANNTR-RRSLTETTTYSHERRTTLTVNGKLSGGGTQFTSSILFG 899
 DB 833 STTINSAYASANNTR-RRSLTETTTYSHERRTTLTVNGKLSGGGTQFTSSILFG 892
 QY 900 KSKKTLSDMAGGYTILSVNTGKPELTLEQTLVYSEKQNPISDLKFTLENDVYDGA 959
 DB 893 KSKKTLSDMAGGYTILSVNTGKPELTLEQTLVYSEKQNPISDLKFTLENDVYDGA 952
 QY 960 LRYLVNDGFEFLNPLTXEQELNDLVYAEQARTLEAKQVEPTAKTQTEGPRYSRA 1019
 DB 953 WRYELVKNGEFLNPLTXEQELNDLVYAEQARTLEAKQVEPTAKTQTEGPRYSRA 1010
 QY 1020 ARAAFPTLLPDQSLNMLEAKQELTAEYQSKA-----KTKYASK----AY 1064
 DB 1011 -RAAFPTLLPDQSLNMLEAKQELTAEYQSKA-----KTKYASK----AY 1069
 QY 1065 FSDPLDQSL-----FA-----LEALVETIAPQOSEKRLAQEAEKQKQK 1108
 DB 1070 LATAVEKETTAQIDPANAQAQLNLTQLEKALAV---AQAQEKERKAQEQAKQKQK 1126
 QY 1109 LISRYNSALSELATVNSMLSVQDELDELFDQAQSAVNTNTAOKRKYDSDAFRAYQ 1168
 DB 1127 LISRYNSALSELATVNSMLSVQDELDELFDQAQSAVNTNTAOKRKYDSDAFRAYQ 1185
 QY 1169 QKTNLQIQVQKALANGELGAVSHSRSDNTDEQVYNNHATLWMSGPAQVOMWDLQEG 1228
 DB 1186 QKTNLQIQVQKALANGELGAVSHSRSDNTDEQVYNNHATLWMSGPAQVOMWDLQEG 1245
 QY 1229 NVGTGASAKMAEESRKIRKALNTGVNASYQFELQQLGQPYGVNRYFERENYQSE 1288
 DB 1246 NVGTGASAKMAEESRKIRKALNTGVNASYQFELQQLGQPYGVNRYFERENYQSE 1305
 QY 1289 EVRYKTPSLAFNRNAGIRVDYTFPTDNISVKEPYFFVNVYDVSNANVGTVNLTVLQOP 1348
 DB 1306 EVRYKTPSLAFNRNAGIRVDYTFPTDNISVKEPYFFVNVYDVSNANVGTVNLTVLQOP 1365
 QY 1349 FGRYWQKEVGLKQELIHFQISAFISKSQSGQSGKQKQNVYGLGYR 1394
 DB 1366 FGRYWQKEVGLKQELIHFQISAFISKSQSGQSGKQKQNVYGLGYR 1411

RESULT 6
 ADA09354 ID ADA09354 standard; protein; 1436 AA.
 XX AC ADA09354;
 XX DT 06-NOV-2003 (first entry)
 XX DE Haemophilus HAP protein from strain 860295.
 XX KW Haemophilus adherence and penetration protein; HAP; antibacterial;
 KW immunostimulant; vaccine; Haemophilus influenza infection; influenza.
 XX OS Haemophilus influenzae; strain 860295.
 XX PN US2003073166-A1.
 XX PD 17-APR-2003.
 XX PF 22-FEB-2002; 2002US-00080505.
 XX PR 25-AUG-1994; 94US-00296791.
 XX PR 20-APR-2001; 2001US-00839996.
 XX PA (GENE/) GENE J W S.
 XX PI Gene JWS;

XX WP1; 2003-567306/53.
 DR N-PSB; ADA09353.
 XX New recombinant Haemophilus adhesion and penetration protein, useful as a
 PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
 PT influenzae infections, in generating antibodies for passive immunization.
 PS Claim 2; Fig 21; 149pp; English.

The invention relates to a new recombinant Haemophilus adhesion and
 penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
 ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
 ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
 acid that will hybridize under stringency conditions to the nucleic acid
 detailed above, an expression vector comprising transcriptional and
 translation regulatory nucleic acid operably linked to the HAP nucleic
 acid, producing HAP by culturing a host cell transformed with the
 expression vector, a composition comprising a pharmaceutical carrier and
 HAP, inducing an immune response in a patient by administering the
 composition, a composition comprising an antibody capable of binding to
 HAP (or to a peptide selected from 35 peptide sequences appearing as
 ADA09359-ADA09393), and an antibody capable of binding to the peptides.
 The Haemophilus adhesion and penetration (HAP) protein is useful as a
 vaccine for prophylactic or therapeutic treatment of an H. influenzae
 infection (e.g. influenza), in generating an immune response in a
 patient, and in generating antibodies to the HAP protein. The antibodies
 may be used to diagnose the presence of H. influenzae infection in a
 sample or patient, in passive immunisation treatments, and to purify or
 separate HAP proteins or the H. influenzae organism from a sample. The
 present sequence is the HAP protein from a strain of H. influenzae.

Sequence 1436 AA:
 Query Match 79.8%; Score 5807; DB 6; Length 1436;
 Beat Local Similarity 79.3%; Pred. No. 0;
 Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;

QY 1 MKKTYFRINFLTACISLIGVSOAMAGHTYRGTQYRPFANKKFTVGAQNTIKYNNQ 60
 DB 1 MKKTYFRINFLTACISLIGVSOAMAGHTYRGTQYRPFANKKFTVGAQNTIKYNNQ 63
 QY 61 GOLVGSMTKAPMIDSVSRNGVALYENQYIVSAHVGVYTDVDCAGANPPDOHRT 120
 DB 61 GLVGSMTKAPMIDSVSRNGVALYENQYIVSAHVGVYTDVDCAGANPPDOHRT 120
 QY 121 YKIKRNNYKKXNDLHPEYDDYHNPRLAKFVTEAPIDMTSNNGSTYSDRTKYPEVRIG 180
 DB 121 YKIKRNNYKKXNDLHPEYDDYHNPRLAKFVTEAPIDMTSNNGSTYSDRTKYPEVRIG 180
 QY 181 SGRQFMRNDQ-----KQDVAGAYHYITLAGNTHNRQAGNGXYSLGSDYRKAQGEYGR 235
 DB 181 TGEQWFRTEDEQXQSKSXLADAVLWRIAGNTHSOSGAGNNTVNLSDITKNNNGP 240
 QY 236 IAGSGDSGSPMTYDAKQKWLINGILRBGNPEKENGFPQVAKSYF-DEIFERDLHT 294
 DB 241 TGVSFDSGSPMTYDAIKQKWLINGVLTGNPESAGNGFQILIKMNFYVDELDPI 300
 QY 295 SLTYRAGNVYTIISGNDNGSGSITQ---KSGIPEIKITLANNSLPLKEXQVHNPRY-- 349
 DB 301 TFIPEPSNGHVSFTSNNGTGVTVQTNNEKVSMPQFKVFTVQ-FNBLAKKK--ZEVYAA 358
 QY 350 DEPNITSPRLANGETTYFNDQKQSLIPASDINGAGGLYFEGNTTYSNNOIOWGAGI 409
 DB 359 GGVNAYKPELNNKNIYFEGDRGTGLITLNNINGAGGLYFEGNTTYSNNOIOWGAGI 418
 QY 410 HVSNGSTVWKVNGVHEDELKSKIGGLTAVQAKENGSGISVQDGKVLLEQADDOGNQ 469
 DB 419 HVEDSTVWKVNGVHEDELKSKIGGLTAVQAKENGSGISVQDGKVLLEQADDOGNQ 478
 QY 470 AFSEIGLVSGRTYQINDDKQPDITKFFYFGRGGRDLNLSHLTKRIQNTDSAMITVH 529
 DB 479 AFKEVGVISGRATYQINSADQVDPPNNIYFGRGGRDLNLSHLTKRIQNTDSAMITVH 538

```

QY 530 NTCQANVTITGNESTVLP--NGNNINKLVREKIAVNGMFGEDTKNKGRLNTIYPTT 588
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 539 NTVANVTITITGNESTVAPSKNNINKLDYSEIAYKMGFEEDENKNGRLNTIYPTT 588
QY 589 EDRTLIISGNTLKGDIOTQTKLFPSPGRTPHAYNHLNKRWSMEGIPQGEIYWDHWI 648
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 599 EDRTLIISGNTLKNKITQEGGTLVFGSRTPHAYNHLN--RPNEL--GRPGSEVVIDDWI 656
QY 649 NRTKANPQIKGGSAYVSRVSSIEGNVTGSNNANAFRGVPRQONTICTESMTGLTT 708
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 657 TRTPKANPQIKGGSAYVSRVSSIEGNVTGSNNANAFRGVPRQONTICTESMTGLTT 716
QY 709 COXVLDPTKVINSIPKQIINGSINTLTDNATANVGLAKLNGVTLTNESQPTLNNATQ 768
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 717 CKYVDLDTKVINSIPTQINGSINTLTDNATVHGLAKLNGVTLTNESQPTLNNATQ 776
QY 769 IGNIRLSNDSTATVDMNLNGVHLTDSAQFSIKNSHPSHQIQGDKGTTVLENNATW 828
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 777 TGNITQLSNHNATVDNANLNGVHLTDSAQFSIKNSHPSHQIQGDKGTTVLENNATW 836
QY 829 SDPTLQMLTNNSTITLNSAYSASSNNTTP--RESLETPTPSHREPTLTVNGKLSGQ 887
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 837 SDATLQMLTNNSTVTLNSAYSASSNNAPEHRSLETTPTSHREPTVTLVNGKLSGQ 896
QY 888 GTPOFTSLFGYKSDKLKLSNDADGYLISVNTGKEBETLEQTLVESKDNQPLSDLKX 947
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 897 GTPOFTSLFGYKSDKLKLSNDADGYLISVNTGKEBETLEQTLVESKDNQPLSDLKX 956
QY 946 FTLENDVDAALRYLKYKNDGEFRLNPIKEQELHNDLVABQAEERTLBAQVEFTNKT 1007
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 957 FTLENDVDAALRYLYLVNNGEERLNPKEQELRNDLVABQAEERTLEAQVEQTAT 1016
QY 1008 QNGEPKRSRRARAFPTLPDSILNLEKQAELEAEPQSKATKYSKRAVSD 1067
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1017 QTSNARVRK--RAVSDTLPDQSQDLVLAQVEPFAEKQNKKA--KKVSKKAAYSD 1071
QY 1068 PLTDOSS-----LFALEALEVID 1085
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1072 TLTPQSQDLVLAQVEPFAEKQNKAKYRSKRAAFPTLPDLSRLKYLEVLAIVIN 1131
QY 1086 APOQSEKRLAQBEBEAKOKKQKDLISRYNSALSISLSTVNSMLSVODELRLPYDQAKS 1145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1132 AQQQYKKE---PODQEKQKQKDLISRYNSALSISLSTVNSMLSVODELRLPYDQAKS 1188
QY 1146 AWMNTIAQDRRRYDSAPRAY--QOKTWRQIGVQKALANGRIGAVFHSRSDNTPEQVK 1205
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1189 AWMNTIAQDRRRYDSAPRAY--QOKTWRQIGVQKALANGRIGAVFHSRSDNTPEQVK 1247
QY 1206 NHATLTWMSGFAYQWGDLOFGVAVGTGISASMAEBOSRKIHKAINTGVNASYQFPLG 1265
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1248 NHATLTWMSGFAYQWGDLOFGVAVGTGISASMAEBOSRKIHKAINTGVNASYQFPLG 1307
QY 1266 QLGIGQPIFYGNRYFIRENYSQSEVRYKTSIAFNRYNAGRIVYTTPTDNTISYKPYFE 1325
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1308 QLGIGQPIFYGNRYFIRENYSQSEVRYKTSIAFNRYNAGRIVYTTPTDNTISYKPYFE 1367
QY 1326 VAVYVVSANVQTTVMVLAQOPGRYMOKEVGLAKETLHFOISAFISQSQSQSLGXON 1385
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1368 VAVYVVSANVQTTVMVLAQOPGRYMOKEVGLAKETLHFOISAFISQSQSQSLGXON 1427
QY 1386 VGVKLGYRM 1394
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1428 VGVKLGYRM 1436

```

RESULT 7
ADA09356
ID ADA09356 standard; protein, 1391 AA.

XX AC ADA09356;

DT 06-NOV-2003 (first entry)

```

XX Haemophilus HAP protein from strain 3219B.
DE
XX
KW Haemophilus adherence and penetration protein; HAP; antibacterial;
KW Immunostimulant; vaccine; Haemophilus influenza infection; influenza.
XX
OS Haemophilus influenzae; strain 3219B.
XX
PN US2003073166-A1.
XX
PD 17-Apr-2003.
XX
PF 22-FEB-2002; 2002US-00080505.
XX
PR 25-AUG-1994; 94US-00296791.
PR 20-APR-2001; 2001US-00839996.
XX
PA (GENE/) GENE J W S.
PI Gene JWS;
XX
XX MPI: 2003-567308/53.
DR N-PEDB; ADA09355.
XX
XX New recombinant Haemophilus adhesion and penetration protein, useful as a
PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
PT influenzae infections, in generating antibodies for passive immunization.
XX
XX Example 2; Fig 23; 14pp; English.
XX
XX The invention relates to a new recombinant Haemophilus adhesion and
CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
CC ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
CC acid that will hybridize under stringency conditions to the nucleic acid
CC detailed above, an expression vector comprising transcriptional and
CC translation regulatory nucleic acid operably linked to the HAP nucleic
CC acid, producing HAP by culturing a host cell transformed with the
CC expression vector, a composition comprising a pharmaceutical carrier and
CC HAP, inducing an immune response in a patient by administering the
CC composition, a composition comprising an antibody capable of binding to
CC HAP (or to a peptide selected from 35 peptide sequences appearing as
CC ADA09353-ADA09357), and an antibody capable of binding to the peptides.
CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
CC infection (e.g., influenza), in generating an immune response in a
CC patient, and in generating antibodies to the HAP protein. The antibodies
CC may be used to diagnose the presence of H. influenzae infection in a
CC sample or patient, in passive immunisation treatments, and to purify or
CC separate HAP proteins or the H. influenzae organism from a sample. The
CC present sequence is the HAP protein from a strain of H. influenzae.
XX
SQ Sequence 1391 AA:
Query Match 72.9%; Score 5303.5; DB 6; Length 1391;
Best Local Similarity 74.8%; Pred. No. 8.1e-297;
Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

```


QY 1 MKETVRLNPLACISLGIVSOAMAGHTYFGIDVOYRDFAEKKGFTYGAQNIKYVNG 60
 DB 1 MKETVRLNPLACISLGIVSOAMAGHTYFGIDVOYRDFAEKKGFTYGAQNIKYVNG 60
 QY 61 GOLVGTSMTKAMIDFSVSRNKGVAALVENQYIVSAHANYGYDVPAGAGNPDQHRPT 120
 DB 61 GEMIGTMKGVMPDLSMVRGCVSTLISEQILISVAHANYGYDVPAGAGNPDQHRPT 120
 QY 121 YKIVKRNNTKKNLHPEDDYHNPRLAKFTYFAPIIDMTSNMGSYSTRKTYERVRIG 180
 DB 121 YKIVKRNNTKKNLHPEDDYHNPRLAKFTYFAPIIDMTSNMGSYSTRKTYERVRIG 180
 QY 181 SCGFQFRNDQK--GDVAGAHYHLLAGNTHNQGAGNGSVYSGDVPAGAEYGLPLTG 238
 DB 181 SCGFQFRNDQK--GDVAGAHYHLLAGNTHNQGAGNGSVYSGDVPAGAEYGLPLTG 238
 QY 239 SKDGSQPMFIYDAEKQKWLINGILREGNPEFGKNGFOLVRKSYEDLIFERDLHTSLAT 298
 DB 239 SKDGSQPMFIYDAEKQKWLINGILREGNPEFGKNGFOLVRKSYEDLIFERDLHTSLAT 298
 QY 299 PAKNGYTTSGDNG----QGSITQSGISEITLTLAMSL--PLAKEDKVAHNPRIQD 352
 DB 299 PAKNGYTTSGDNG----QGSITQSGISEITLTLAMSL--PLAKEDKVAHNPRIQD 352
 QY 353 NIVSPRLNNGETLYFMQKQSLIFASQINQAGSLYFEGNFTVSPNSNQ--TMQAGIHY 411
 DB 353 NIVSPRLNNGETLYFMQKQSLIFASQINQAGSLYFEGNFTVSPNSNQ--TMQAGIHY 411
 QY 412 SENSTVTKVNGVEHDLRLSKIGKTLHVQAKENKGSISVGDKEVILLEGQADQGNQAF 471
 DB 412 SENSTVTKVNGVEHDLRLSKIGKTLHVQAKENKGSISVGDKEVILLEGQADQGNQAF 471
 QY 472 SEIGLVSQKTYQLNDQKQPTDKYFGFRGRLDNLNHSITLTKRIQNTDEAMTHNHT 531
 DB 472 SEIGLVSQKTYQLNDQKQPTDKYFGFRGRLDNLNHSITLTKRIQNTDEAMTHNHT 531
 QY 532 TOAANVTIGNSIVLPNGKNTKLDYRKEIAYNGFGEPIKQKHNKRLNLTLYETEDR 591
 DB 532 TOAANVTIGNSIVLPNGKNTKLDYRKEIAYNGFGEPIKQKHNKRLNLTLYETEDR 591
 QY 592 TLLISGGINLKDITQYTKGLFFSGRPTPHAVNHLNKKWSEMGIPQGEIWMHDWIKRT 651
 DB 592 TLLISGGINLKDITQYTKGLFFSGRPTPHAVNHLNKKWSEMGIPQGEIWMHDWIKRT 651
 QY 652 FKAENPQIKGSAVSRVSSIEGNMTVSNMNAATFGVVPQONCTCTSWTGTITTCOK 711
 DB 652 FKAENPQIKGSAVSRVSSIEGNMTVSNMNAATFGVVPQONCTCTSWTGTITTCOK 711
 QY 712 VDLDTVYNSIPKQINGSLNLTNATATAVAGLAKAGNYTLTNSQFTLSNNATQIGN 771
 DB 712 VDLDTVYNSIPKQINGSLNLTNATATAVAGLAKAGNYTLTNSQFTLSNNATQIGN 771
 QY 772 IRLSDNSTATVDNANLNGVHLLDSAQFSIKNSHPSHOIQGDKGTYVLENNATWPSDT 831
 DB 772 IRLSDNSTATVDNANLNGVHLLDSAQFSIKNSHPSHOIQGDKGTYVLENNATWPSDT 831
 QY 832 TLQNLTLNNTITLNSAYASASNNTP--RPRSLRETTPTSAHREFTLTVNGKLSGQGT 889
 DB 832 TLQNLTLNNTITLNSAYASASNNTP--RPRSLRETTPTSAHREFTLTVNGKLSGQGT 889
 QY 890 POFSSLEFGYKDKLKLSDAEGDYILSVNTKEPETHLEDTLVESKONQPSDKLKEFT 949
 DB 890 POFSSLEFGYKDKLKLSDAEGDYILSVNTKEPETHLEDTLVESKONQPSDKLKEFT 949
 QY 950 LENDHVAGALRYKLVKNDGEFRLNPIKEQELHNDLVYRQAQERTLEAKQVPEPAKQYT 1009
 DB 950 LENDHVAGALRYKLVKNDGEFRLNPIKEQELHNDLVYRQAQERTLEAKQVPEPAKQYT 1009
 QY 1010 GEPKRSRRARARAFDTLPPQSLNLNLEAKQELAEQOKSKKTKKVSRAV--RSD 1067
 DB 1010 GEPKRSRRARARAFDTLPPQSLNLNLEAKQELAEQOKSKKTKKVSRAV--RSD 1067
 QY 1067 SKARVSR--RAVFDPLPAOSLLKLEAKQK--LTTEIYISKA--KVASRRARARERSD 1067
 DB 1067 SKARVSR--RAVFDPLPAOSLLKLEAKQK--LTTEIYISKA--KVASRRARARERSD 1067

QY 1068 PLLDSLFALAEALVIDAPQOSEKRLAQEAEK--QRKQKLLISRYNSALSLSATVN 1126
 DB 1068 TLPDQ---LQALVELIADQQVKKERQTEEEERKQKQKELLISRYNSALSLSATVN 1124
 QY 1127 SMLSVDLRLRVDDQASAVWTNIAQDNRHYSDAFPAVQOQCTNIRQIGYQKALANG 1186
 DB 1127 SMLSVDLRLRVDDQASAVWTNIAQDNRHYSDAFPAVQOQCTNIRQIGYQKALANG 1186
 QY 1187 ICAVFSHSRSDNTPDQVKNHATLTMNSGPAQYQWGDLOFGVNTGTISASMAEBSQRX 1246
 DB 1187 ICAVFSHSRSDNTPDQVKNHATLTMNSGPAQYQWGDLOFGVNTGTISASMAEBSQRX 1246
 QY 1247 IHRKAINGVASAYORLQGLIQPYGVNRYPIERENYQSEBVRVTPSLAENRYAGI 1306
 DB 1247 IHRKAINGVASAYORLQGLIQPYGVNRYPIERENYQSEBVRVTPSLAENRYAGI 1306
 QY 1307 RVDYFTPTDNISYKPYFVNRYVYSNANVQTVNLTVIQPFGRYQKQVGLKELLHF 1366
 DB 1307 RVDYFTPTDNISYKPYFVNRYVYSNANVQTVNLTVIQPFGRYQKQVGLKELLHF 1366
 QY 1367 QISAFISKQSGSLGKQKQWGVKLGYN 1394
 DB 1367 QISAFISKQSGSLGKQKQWGVKLGYN 1394

RESULT 9
 ABU37787
 ID ABU37787 standard; protein; 1449 AA.
 AC ABU37787;
 AC 19-JUN-2003 (first entry)
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #23314.
 XX
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 XX Neisseria meningitidis.
 OS
 XX
 XX MO200277183-A2.
 PD 03-OCT-2002.
 PF 21-MAR-2002; 2002MO-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 PA (ELIT-) BLITRA PHARM INC.
 XX
 XX Wang L, Zarnudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PESDB; ACA41657.
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 25; SEQ ID NO 65711; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding

the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1449 AA:

Query Match 57 48: Score 4178; DB 6; Length 1449;
Best Local Similarity 56.28; Pred. No. 6.7e-232;
Matches 816; Conservative 214; Mismatches 325; Indels 98; Gaps 17;

10 FLTACISLIGVQAAGHTFEGJIDYRDFENKGFETVGAQNIKYNKQOLVGTST 63
27 YLALCISFELLPGMAKGTTFGLNYRDFENKGFETVGAQNIKYNKQOLVGTST 86
70 KAPMIDFSVSNVGAALVENVYIVSAVAVGCTVDVPAEKNPDPQHFYTKYKANN 129
87 KAPMIDFSVSNVGAALVENVYIVSAVAVGCTVDVPAEKNPDPQHFYTKYKANN 146
130 KCONLHPYEDDTHNRLKFTTAAPIDMTSMNCSYSTRTPYRPTSGSGQPRND 189
147 KPDMSHPYGDYHMLKFTTAAPIDMTSMNCSYSTRTPYRPTSGSGQPRND 206
190 QDKGDVAGAVHYLTAGNTNORAGNGSYSGVVRKAGEYPLPIAGSKDSGSMF 249
207 DDKRGLSYSGAMLCGTHMGKANNVYSLSDVHRANDYPMPIAAGDSSSMF 266
250 YDAERKQMLNGLBEANPEFGKNGFQIVKSYF-DEIFERDLATSLYTRAGVYTIS 308
267 YDKNKMLNGLBEANPEFGKNGFQIVKSYF-DEIFERDLATSLYTRAGVYTIS 326
309 GNDGCGSTTQSGIPSEIKITLANNST--PLXKDKVNHPRY--DGNITYSPRLNNGE 363
327 SNNNGTGTVEETNEKYNPKLVYALPDPSINETDK--EPYIAGAGVNOYRRRLNNGE 384
364 TLVEMQKQKSLIFASDINOGAGLYFEGNFTVSPNSQTMQAGIIVSENSTVTVKNG 423
385 NLSFIIDYNGKILNININOGAGLYFEGNFTVSPENNETMQAGVHISDSTVTVKNG 444
424 VEHDRLSKIGKGLTHVQKNGKSSISVDQAVITLEQADQKQKAFSEIIVSNGRTV 483
445 VANDRLSKIGKGLTHVQKNGKSSISVDQAVITLEQADQKQKAFSEIIVSNGRTV 504
484 QLNDKQFOTDKFYFSGRGLDLNGHSLFKRIONTBEGMIVNNTTQAANVTITGNE 543
505 QLNDKQFOTDKFYFSGRGLDLNGHSLFKRIONTBEGMIVNNTTQAANVTITGNE 564
544 SILVPGNNINIKLDYRKETIANGWFGETDNKHNRLNIYKPTTDRITLLSGGTLNKG 603
565 SILVPGNNINIRNLNYSKELIANGWFGKOTTKNRLNLVYQPAEDRTLLSGGTLNKG 624
604 DITVTKKLFSGRPLPPIANVNLKMSKSEKIGIPOGLVWHOMIKNTKANFQIKSGS 663
625 NITVTKKLFSGRPLPPIANVNLKMSKSEKIGIPOGLVWHOMIKNTKANFQIKSGS 684

664 AVASRNVSISGNMTVSNNAATGVVNOQNTICRSMTGLTTCQKVDLDTYVINSI 723
685 AINSNAKVEGOMHLSNNAQAVGVAPHQSHITICRSDMTGLTTCQKVDLDTYVINSI 744
724 PPTQINSIMLTNATANKVLAKNNTVLRHSQFLLSNNAQVGN 771
745 TTDLSGNVSLADHNLTLATLNGNSANGDRYVSNATQNSLVLGNNAQATFN 804
772 -----IRLSNSTATVDNNLNLNGVHLTDSNOFLSNS 804
805 QATLNGTASGNASVNSNBAQSGSLTSDAKANVSHALANGVSLADKAVHFNENS 864
805 HSHQIQDQKGTITLENATWPSDTTLQNTLNNSTITNSAV-----SASSNNT 856
865 RETQLSGSDMTALHKQSEWTLPSGTGLNMLDVATITNSAVRHDAQAQTSVSDT 924
857 PR--RSLTETTPSAHRNTLITVNGLSQGFQFQTSLSYKSKDKLSNABSD 913
925 PRRSRRLSLAVPPPTVSRENTLITVNGKNGQGFRRFMSLQYSDKLSKABSSBST 984
914 YILSVRNTGKEPHTLEQLTVESKDNQPLSDKLFLENDVAGALRYKLVNDGEFRL 973
985 YTLAVNNTGNEPVSLQVLTVDGKNKPLSNNLFTLQNHVDAGAMRYQLIKQSEFRL 1044
974 HNPKEQELHNDLVRAE---QAE-----TEAQVEPTAKTQT-GEPRVSRPARA 1022
1045 HNPVEQELSDKLGKAEKQAEKQKDNQASIDALIAAGRDAAKTESVAEP-----ARQ 1097
1023 APPDTLPQSLNNLNAKQAEITLAEYKSKAKTKKRSKRAVSPDPLDQSLALEMALE 1082
1098 AGEVAV--GIMQAEKEX-RVQADKQSAKQREKRETPATVFP-----BARRA 1144
1083 VIDAPQSEKORLAQEAQKQ-RKQKOLISRYNSALSELATVNSMLSYODELDLFPV 1141
1145 RRLDFQ-----PQPOPQPOPQRLISRYNSALSELATVNSVFNQODELDLFPV 1197
1142 QASAVMTNIAQDKRRYSDAFRAYQOQKTNLPQIGVQALANGRIGAVSHRSRSDNTFD 1201
1198 DRNNVMTSGRDTYHRSQDFRAYRQQ-TDLRQIGMGNLSGSGVGLFISHRNTENTD 1256
1202 EYQKHAITLTMGSPAYQYQGLQFQVNVGVSISAKMAESQSKIKHAKINQVNASQ 1261
1257 DQIGNSARLASHAVFVGQYIGRFDIGISTGAGSFGSLDQIGGIRRVRLVHGIQARR 1316
1262 FRIGQLGIOPYGVNRPTEIRENOSSEVRKYTPSLAFNRVAGIRVDYPTPTDNISYK 1321
1317 AGFGGGLPEYTGATRYFQKADRYEVNINATPGALFNRVAGIRVDYPTPTDNISYK 1376
1322 PYFPVNVYDVSNANVQTTNLTVLQOPGRYQKQVGLAETLHFOJISAFISKSQSGQLG 1381
1377 PYLSSTYDASGKVRKRYNTVLAQDFGKTRSAEMGVNAEIKGFTLISHAAAKGPOLE 1436
1382 KQNVGVKLSYRW 1394
1437 AQSAGIKLSYRW 1449

RESULT 10

AAV56621 standard; protein; 1454 AA.

AAV56621;

21-FEB-2000 (first entry)

Neisseria meningitidis strain ATCC13090 BASB006 protein sequence.

Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;

KM vaccine; antibiotic; upper respiratory tract infection; meningitis;

KM invasive bacterial disease; bacteraemia; screening; antibacterial.

OS Neisseria meningitidis.

Y	1379	QLGKOONVYGLGRM	1394
		: :	
	1439	QLBAQHSKIGLGRM	1454
Db			
	RESULT 11		
	AA58592		
	AA58592 standard; protein; 1457 AA.		
XX	AA58592;		
AC			
DT	13-MAR-2001 (first entry)		
DE	N. meningitidis amino acid sequence orfl-1 pep SRQ ID NO-87.		
XX			
XX	Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;		
KM	diagnosis; antigen; detection; infection; gene therapy; antibacterial.		
XX			
OS	Neisseria meningitidis.		
XX			
XX	WO20006791-A1.		
PN			
XX	09-NOV-2000.		
PF	08-MAR-2000; 2000MO-US005928.		
XX			
XX	30-APR-1999; 99US-0112068P.		
PR	08-OCT-1999; 99MO-US023573.		
PR	28-FEB-2000; 2000GB-00004695.		
XX			
XX	(CHR.) CHIRON CORP.		
PA	(GENO-) INST GENOMIC RES.		
PI	Pizza M, Hickey E, Peterson J, Tettein H, Venter JC,		
PI	Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,		
PI	Rappuoli R, Frazer CM, Grandi G;		
DR	WPI; 2000-647603/62.		
DR	N-PDB; AAF21587.		
XX			
XX	Neisseria meningitidis B full length genome sequence and open reading		
PT	frames are used to detect, treat and prevent Neisserial infections.		
XX			
XX	Example 1; Page 114; 692p; English.		
XX			
XX	The present invention describes the full length genome of Neisseria		
CC	meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613		
CC	represent fragments of the NMB genomic sequence, as the sequence was too		
CC	long to go in a record on its own it was split into 8 sequences by 4980 bp		
CC	overlap each other at the beginning and end of each sequence by 4980 bp		
CC	(i.e. the last 49980 bp of AAF21544 is repeated at the beginning of		
CC	AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of		
CC	AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins		
CC	given in AA58592 to AA58593, and AAF21589 to AAF21606 represent PCR-		
CC	primers which are used in the exemplification of the present invention.		
CC	The NMB genome and fragments from it have antibacterial activity, and can		
CC	be used in vaccines and gene therapy. Neisseria nucleic acids, proteins		
CC	and/or antibodies which binds to the proteins can be used in compositions		
CC	for treating or preventing infection due to Neisserial bacteria or as a		
CC	diagnostic reagent for detecting the presence of Neisserial bacteria or		
CC	of antibodies raised to Neisserial bacteria. Computers, computer memory,		
CC	computer storage medium or computer databases can be used in a search to		
CC	identify open reading frames (ORFs), or coding sequences within the NMB		
CC	genome. The DNA sequences provide further opportunities to find antigenic		
CC	or immunogenic proteins which are more effective in vaccines		
XX			
XX	Sequence 1457 AA;		
SQ			
	Query Match	56.2%; Score 4085; DB 3; Length 1457;	
	Best Local Similarity	55.1%; Pred. No. 1.ee-226;	
	Matches	806; Conservative 219; Mismatches 328; Indels 110; Gaps 19;	
QY	10	FLTKSLGIVGQAMAHGHTYFGIDVYRRDPAENKRGKFTWGAQNIKTVKKGQGLVGRSM	69

Dh	27	YLAICLSFGILQAAWAGHTVGLNVOYTRDPELNKGFANGAODIEVYNNKGEIVGMSMT	86
Qy	70	KAPIDPVSYSNNGAALAVENQYIVSAHNNGYTDVDFGABGNPDORFYTKIVKNNY	129
Dh	87	KAPIDPVSYSNNGAALVABQYIVSAHNNGYNNVDFGABGNPDORFYTKIVKNNY	146
Qy	130	KKCNL-HPEYEDYHNPRLHKEFTAEAPIDMTSNNGSTYSDRTKYERPRIGSSORWEN	188
Dh	147	KAGTKGHPYCGDYENHPRHKEFTVDAEPVMTSYNDGRYTYIDONNYPDRYIAGAGYWRSS	206
Qy	189	DODGCD-----QVAGAHYCTAGNHNORGAGNYSVLGADVRKAGEGPLTAGSKGD	242
Dh	207	DEDEPNNRSSYHIASASWLVGNTPAONGSGGTYVNLSEKIKASPGFLPTGSSFGD	266
Qy	243	SGSEMPFYDAEKOKVLINGILREGENPEGEKGFOLYKSYF-DEIFERDHTLSYTRAG	301
Dh	267	SGSEMPFYDAQOKVLINGVLQGNPNYIKSXNGFOLYKMPFYDEIFAGDHSVYEBRQ	326
Qy	302	NGVYITSGNDNGQSSITOK---SGISEIK---ITLAMSILPKEXDKYHNRYDGNPY	355
Dh	327	NGKTSFPMDDNNGGKIKNAHEHNSLPRLKRTYVLFVSLSTBAEPYHNA-GENYST	385
Qy	356	SPRLNNGETLYPMDDKOSLLFASDINQAGGLYEBGNFTYSPNSQOTQAGIHSVENS	415
Dh	386	RPRLNNGHNSIFDEKKEHLLTSNINQAGGLYQSGO?YVSPNNHTQAGVHISDS	445
Qy	416	ITYWKNGVNEHDLSTKIGKGLIHAVQAKENKSSIVGDCGYILLFQADDOGNKQAFSSIG	475
Dh	446	ITYWKNGVNANGELSTKIGKGLIHAVQAKENQSSIVGCGVILLDQOADDKGKQAFSSIG	505
Qy	476	LYSRRGYQVANDKQDPDIFYGFBGRGLDINHSILPERIONTDEGAMTYNHTTQAA	535
Dh	506	LYSRRGYVQVANDKQNPDKUYGFBGRGLDINHSLSFPHRIQNTDEGAMTYNHNQDES	565
Qy	536	NVTTIGNESTVLEPNCNNIKLDYKELIAYNGWFGETDNKNKINGRLNYKPTTEDRTLL	595
Dh	566	YTTITGNKMDLA-ITGNK-NSLDSKKEIAYNGWFGEDTTKNGRLNLYQPAEDRTLL	623
Qy	596	SGGTNLKODITOTYKGLFFSGRPPTPHAYNHLNKRSEWBGIDOEIYOMHDYNTFPAE	655
Dh	624	SGGTNLKMNITOTNGGLFFSGRPPTPHAYNHLNDHNSQKESIPGEIYOMHDYNTFPAE	683
Qy	656	NFOIKSGAAVANSNYSIEGMWTVSNNAKATFCVVPNOQNTICRSMDGLTYTCOKVLT	715
Dh	684	NFOIKSGAAVANSRKYAKVGDHLSNNAQVBPVAPHOSHITCRSDMTGLNVEKILT	743
Qy	716	DTKVLINSIPKTOINGSINLTUNA?ANYKGLAKINXVLTINHSOFTLSNNAQIDN----	771
Dh	744	DDKYIASLTKTIDISGNVDLADHMLNLTGLATINGNLSANDGTITYVSHNAQTNGSLTY	803
Qy	772	-----IRLSDNSTATVONAMNGVHLLDS	796
Dh	804	GNAGQTFNQAALINGNTSASGNASPYLSDHAVONGSLTISGNAKANVSHSLMKNYSLADK	863
Qy	797	AOFSLSNHSFHOQDCKGTVLTLENAITWMPSDTTLQNTLANSSTLTNSAY-----	849
Dh	864	AYPHHESRFLTCQISGGXOTJALHKDSBMTLPBGTELGNMLDNATITLINSAYRHDAGA	923
Qy	850	-SASSNNTR-----RSLSETTTPISAEHRENTLTYNGKLSGCGEFOPTSLFEGYSD	902
Dh	924	QTSBATTDAPRRSHRSRRSLSVTPPTSVBSRNYTLTVGNKINQGFPRMSLFPSTSD	983
Qy	903	KKLNSDAEGYIISVANNNGKEPTEGLTIVSSKXNQSIDKLETTENDHDVADGALRY	962
Dh	984	KKLKASSEGYTLAANNKTRPABLEDLYVSGKDKRFLSNLNFLLQNEHVDAQWRY	1043
Qy	963	KLVKNDGEFRLNFTKEOELANDIVRAE---QABR-----TLEAKVEPTKQOT-GE	1011
Dh	1044	QILRKQGEFRLNFPKEOELSDKLGKAKKQAKQEKMAOSIDALITAGRAVETKESVAE	1103
Qy	1012	PKRKSRAAARAPDPLPOOSILNLAKQAEITATQSKAKTKKVSRAAFSPDILD	1071

```

Dh      804 GNAATATVQATLAKNTSASGNASPNLSHVAQNGSLTSLGNKAVNSHSLANVSLADK 863
Qy      797 ACPSTLKNHFSHQIQGKGTVTLENACTWPSPTTLQNTLNNTSTILNSAY----- 849
Db      864 AVHFHSSSPFTGQISGKDTALHKKOSEWTLPSGTGLGAMLDNNTITLNSAHRHADA 923
Qy      850 -SASNNVTR-----RRSLEETPTSAHRFNTLTWNGKLSGGGTFOPTSSLEFGKSD 902
Db      924 QTSATADAPRRRRSRRLSLSTPPTSVESRFTLTWNGKLNKGOTRFWSELEFGHSD 983
Qy      903 KKLKSLNDAGDTLLSVNTEKEPETLEGLTVESKMDQPSDKLFTLENDHVDAGALRY 962
Db      984 KKLKSLSSSGTTLAVNNTGNBPASLEQLTVESGDKKPLSEKINFTLQNEHVAGAMRI 1043
Qy      963 KLVNNGEFLRNPKEQELHNDLVAR---QAEK-----TLKQVEPTACTQY-GE 1011
Db      1044 QLRKQSEFLRNPYKQESLSDKGLAEKKQAEKDAOSLDLILAAGRDAVEKTSVME 1103
Qy      1012 PKVSRRAARAFPTLTPDOSLNLALAYQAELETFQSKAKTKKVSKAAPVSDPLD 1071
Db      1104 P-----ARQAGEHV---GIMQAEERK-RVQADDTALAKQREARREATTAF- 1149
Qy      1072 QSLFALAEVITDAPQOSEKDLAQBELKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 1131
Db      1150 -----BARARRDLPQLQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 1195
Qy      1132 QDELRLFLVDAQASAVNTINADQRRYSDAFRAVQOQKTNLQIGVQALANGRIGAVT 1191
Db      1196 QDELDRVFAEDRRANVWTSIGRTKTRHSDQDRAVQ-QIDLQIQMKNKNGSRVGILF 1254
Qy      1192 SHSRSDTPEQVQKHATLTWMSGFAQYQMDLQGVNVGTGISAKMAEEOGRKHKA 1251
Db      1255 SHNRTXTFDDGIGNSATLNGAVFGQYQIDRFYIGISAGKSSSLDGLGKIRRRY 1314
Qy      1232 INGVWASVQFRLQGLQYFYGNRRFIERENYQSEVRYKPSLAFNRVAGIRVYT 1311
Db      1315 LHYGIQARYDAGFGGGLBPHIGATRYFGKADYEVENVNATPGLAFNRVAGIKADYS 1374
Qy      1312 FTPTDNTSVKPYFVYVYVSNANVQTTNVLTVLQOPFGRYQKQVGLAEETLHPQISAF 1371
Db      1375 FEPKQHISTFTYLSSTYDASGKVRTRVNTVNLADFGKTRSMWGNALKGPITLSH 1434
Qy      1372 ISKQSGSQLKQONVVKLGRW 1394
Db      1435 AAAAGPQLBAQSHAGIKLGRW 1457

RESULT 14
AAVS6622
ID      AAVS6622 standard; protein; 1457 AA.
AC      AAVS6622;
XX      21-FEB-2000 (first entry)
Dc      Neisseria meningitidis strain H44/76 BASB006 protein sequence.
XX      Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
KM      vaccine; antibiotic; upper respiratory tract infection; meningitis;
XX      invasive bacterial disease; bacteraemia; screening; antibacterial.
XX      Neisseria meningitidis.
OS      Neisseria meningitidis.
PN      W0995873-A2.
XX      04-NOV-1999.
Dc      20-APR-1999; 99W0-BP002766.
XX      24-APR-1998; 98GB-00008866.
PA      (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

```

```

XX      Thomas D;
XX      WPI; 2000-052810/04.
DR      N-PSDB; AA38918.
XX      Novel polynucleotides and polypeptides from Neisseria meningitidis used to
PT      prepare vaccines against bacterial infections.
PS      Claim 3; Page 93-98; 103pp; English.
XX

```

The present sequence is BASB006 isolated from *Neisseria meningitidis* strain H44/76. BASB006 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for disease, particularly human diseases. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridisation probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutation in BASB003 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as upper respiratory tract infection, invasive bacterial diseases such as bacteraemia and meningitis, and for the development and screening of antibacterial drugs. They are also used in the prevention of adhesion of bacteria to eukaryotic matrix proteins on in-dwelling devices, or to extracellular proteins on wounds, and to prevent tissue damage and/or block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques

Sequence 1457 AA:

Query Match 56.1%; Score 4083; DB 3; Length 1457;
 Best Local Similarity 55.0%; Pred. No. 2e-226;
 Matches 805; Conservative 220; Mismatches 328; Indels 110; Gaps 19;

```

Qy      10 ELTACTSIGIVSQMAHGYTFGIDGYRDFRANKKFLTVGAKNTKYVKKQGLVTSMT 69
Db      27 YLAICLSFGILPQMAHGYTFGINTQYRDFRANKKFLTVGAKNTKYVKKQGLVTSMT 86
Qy      70 KAPMIDFVSNSGVNVAIVEMQYIVSAHVNGYTDVDFGAGNBPQHRFTYKLVKNNY 129
Db      87 KAPMIDFVSNSGVNVAIVEMQYIVSAHVNGYTDVDFGAGNBPQHRFTYKLVKNNY 145
Qy      130 KKQNLHPYEDDYNPRLAKFTYTAAPIDTSMNNGSTISDRTPKRYRIGSGRQFERN 188
Db      147 KAGTKCHPYGSDYHNPRLAKFTYTAAPIDTSMNNGSTISDRTPKRYRIGSGRQFERN 206
Qy      189 DQDKGD-----QVAGVHYLTGNTHNQRGANGSYLGQGVYKAGRGYGLPIASKGD 242
Db      207 DDEPPNRSYHIAASVWLVGNTFPQNSGGTYNLSSEKIKHSPYGLPFGSGFD 266
Qy      243 SGSPMYFYDAEKQKMLNGILIREGNPFEGKNGFQVLRKSYF-DEIFERDHTSLYTRAG 301
Db      267 SGSPMYFYDAEKQKMLNGILIREGNPFEGKNGFQVLRKSYF-DEIFERDHTSLYTRAG 326
Qy      302 NGVYITSGDNQGSSTQK--SGITSEIK--ITLAMSLLPKEDKKNHPRRDGNIT 355
Db      327 NGKYSFVNDNNGTGKINKAKHNSLNPRLKTRTYQLNVSLSETPAEPIYHAA-GGVNNT 385
Qy      356 SPRLNNGETLYMPQKQSLIFASDINQAGGLYEGNGFTVSNSTQVQAGIIVSBN 415
Db      386 RPRLNNGENISPTDGEKELLISNINQAGGLYEGNGFTVSNSTQVQAGIIVSBN 445
Qy      416 TVTAKVGYEHDRLSKIGKTLHYQAKGENKGISVGDGVYILLEQALDQGNKQAFSEIG 475
Db      446 TVTAKVGYEHDRLSKIGKTLHYQAKGENKGISVGDGVYILLEQALDQGNKQAFSEIG 505
Qy      476 LVSGRGTVQLNDKQPTDCKYFGRFGSLDINSHSLTFKRIQNTQEGAMINHTTQAA 535

```

```

Db 506 LVSGRTVQVQALNADQNPDIYFGRGRLDLNGLSHSPHRIQNTDEBGMVNHNDOKES 565
Qy 536 NNTTNGNESIVLPNGNNINKLDYRKELIAYNGPGETDKNGNGBRLNLYKPTEDRTLL 595
Db 566 TTTTGNKQIA-TTGN-NSLDSKKEIAYNGPGBKTTTKRKRLNLYQPAARDRIILL 623
Qy 596 SGGTNKSGDITQTKGLFFSGRPTPAVYHNLKWSMEBGIPOGEIWDHDMINRTPKAE 655
Db 624 SGGTNLNGNITQTKGLFFSGRPTPAVYHNLNDHMSQKGIPIGELIWDNDMINRTPKAE 683
Qy 656 NFOIKGSAVSRNVSSIEBNTVSNANATFVVPVQNTITCSDMTGTITQAKDUL 715
Db 684 NFOIKGSAVSRNVAKYKDMHLSHQAQVGAAPQSHITCTSDMTGLTNCVEKTTI 743
Qy 716 DTFVINSIPKQIOMESINTLNDNAFVWGLAALNGVTLTNHSGFTLSNNAQTGN---- 771
Db 744 DDKVIASLTXTDISGVNDLADHMLNLGLATLNGNSANGDTRIVSHNATQNGNLSLV 803
Qy 772 -----ILSDNSTATVYDANLNGVHLTDS 796
Db 804 GNAQATPQATLNGNTSASGNASPNLSDAVQNSLILSGNAQANVSHALNGVSLADK 863
Qy 797 AGSLKNSHPHQIOGDKTTVLENAWTMPSDTTQNTLANSITTLNAY----- 849
Db 864 AVHFSSSRFTQISGKDTALHLKQSEWTLPSGTELGNTLNLNATITLNSAYHDAAG 923
Qy 850 -SASNNTR-----RSLETETPTSAHFNTLTNGKLSGGTFOFTSSLFGYKD 992
Db 924 QTSACDAPRRSRSRSLSTPPTSVESFNITLTNGKLNQGTFRFSELPGRSD 993
Qy 903 KKLKSDAAGDYILSVNTGKEPTELEUTLIVESKQNPISDKLKTLENDHVDAGALRY 962
Db 984 KKLKASSSGTYTLAVNNTGNEPASLEQTLVVEGKNKPLSENFNTLLNEHVDAQAMRY 1043
Qy 963 KLVNDGEPRLNPIKEDELANLVRAE---QAEP-----TLBAQVEPACTQY-GE 1011
Db 1044 QLIKQGEFRLNPIYEOELSDKLGAEKKAQAEKQNAOSIDMLIAGDAVETKESVAB 1103
Qy 1012 PKVSRRAAARAPDTLPDQSLNALAEAKOALTFETQSKAKTKVRSKRAVSDPLLD 1371
Db 1104 P-----AKQKGENV---GIMQABEKK-RVQDKXTALAKQRAETRPRTIAFP--- 1149
Qy 1072 QSLFALAELEVIDAPQOSEKRLAEBEKQKQKDIISRYNSALSLSATVNSMLSV 1131
Db 1150 -----PARARARDLPQLOPQOPQORDIISRYNSGLSSEFATLNSVFAV 1195
Qy 1132 QDELDRLVYDQASAVNTNIAQDKRISDAPRAYQOQKTMLOIGVQALANGRIQAVF 1191
Db 1196 QDELDRLVFAEERNNVMTSGIRDTKHRSODFRARQO-TLHQIQMOKNLSSGRVGLF 1254
Qy 1192 SHSRSDTPEQVQKHATLTMWSGFAQYQMGLOGFVWGTGSASMAEBQSRKIRKA 1251
Db 1255 SHNKTENTPDGIGNSALAGAVFGQYIGDFYIGISAGAGSSGSLDGGKIRRV 1314
Qy 1252 INGVNNSYQFLQGIQPYFGVNRFFIRSNYSSEVRVKTPSLAFNRNAGIRVDYT 1311
Db 1315 LHYGIQARYAGGCGGLEPHIGATRYFOKADYREYVWJNATFGALFRYRAGIKADYS 1374
Qy 1312 FTETDNTSVKPYFVWYVDSNANVQTTNLTLYLQPPGRWQKVEGLAAELHQSASF 1371
Db 1375 FKPAQHSITPPLSYLTDASGCVTRVNTAVLADQFGKRSLEWMAHETKGTLSLH 1434
Qy 1372 ISKSQSQLGKQONVGVKLYRM 1394
Db 1435 AAAAKGQOLAGHSAGIKLGYRM 1457

```

RESULT 15
 AAB25662
 ID AAB25662 standard; protein: 1457 AA.
 AC AAB25662;
 XX

```

DT 04-DEC-2000 (first entry)
XX
D2 N. meningitidis amino acid sequence orfl-1.pep SEQ ID NO:1047.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
XX antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB.
XX
OS Neisseria meningitidis.
XX
PN MO200322430-A2.
XX
PD 20-APR-2000.
XX
PE 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-003794P.
XX
PR 30-APR-1999; 99US-0132068P.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masiagnani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
DR N-PSDB; AAB1302.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX
XX
XX Example 1; Page 113; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAB81453 to AAB82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAB81260 to AAB81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAB81254 to AAB81259 and
CC AAB81304 to AAB81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAB81322 to AAB81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 1457 AA;

```

Query Match 56.14; Score 4082; DB 3; Length 1457;
 Best Local Similarity 55.14; Pred. No. 2,36-226;
 Matches 806; Conservative 218; Mismatches 329; Indels 110; Gaps 19;

```

Qy 10 ELKAGISGIYQAMAGHTFEGIDVOYRDFEAKKGFYGAONIKVYNNKQGLVGTSMT 69
Db 27 YLAICLSFGILPQMAAGHTTFGINTOYTRDPAENKKGFAVGAKDILFYNNKGELVGSMT 86
Qy 70 KAPMIDFVSVNSNGAALVENQYIYVVAHNVGYTDVDFGAEAGNPDQRFYTKIVKKNY 129
Db 87 KAPMIDFVSVNSNGAALVGDQYIYVVAHNGYVNVDFGAEGRNPDQRFYTKIVKKNY 146

```

QY 130 KCONL-HPEDDYHNPRLKTEAAPIDMTSNMGSYSDDRTKYERVRIGSGRQPMR 188
 Db 147 KAGKRGHPYGSYHNPRLKTEAAPIDMTSNMGSYSDDRTKYERVRIGSGRQPMR 206
 QY 189 DODKGD-----QVAGAVHYIAGNTENKRGANGSYISGQDVRKAGEXYPLIAGSKD 242
 Db 207 DEDEPNRRESSYHIASYSMLVGNTPAQNCGGGTVALGSEKIKHSPYGLPFGSSFG 266
 QY 233 SSGMFTYABEKOKMLINGILREGNPEKKNRQVLRKYF-DEFEEDLHSTLYRAG 301
 Db 267 SSGMFTYABEKOKMLINGILREGNPEKKNRQVLRKYF-DEFEEDLHSTLYRAG 326
 QY 302 NGVYISGNDNGOQSITOK---SGIDSEIK---ITLAMSILPKKOKKXNPRDGNIT 355
 Db 327 NGKYSRDDNNGTGKINAKHEHNSLPRKTKTYOVLPRVSLSTAEPEYHAA-CGNVNS 385
 QY 356 SPRLNWGETLYPMDOQKOSILI PASDINGAGGLYEBGNFTVSPNSQWQAGIHSVENS 415
 Db 386 RPRLNWGNEMISFIDEGKGLITNSINOGAGGLYFQSDFTVSPENNETHWQAGVHISDS 445
 QY 416 TVTKVNGVYHNRSLKIKGKTLHQAAGNKGSI SVBGKVTILEQADNQGKAPSEIG 475
 Db 446 TVTKVNGVANDRLSKIKGKTLHQAAGNKGSI SVBGKVTILEQADNQGKAPSEIG 505
 QY 476 LVSGRGVQINDKQFOTDKFYGFGRLDLNGHSLTFKRIQNTDEGAMIVNHTTQAA 535
 Db 506 LVSGRGVQINDKQFOTDKFYGFGRLDLNGHSLTFKRIQNTDEGAMIVNHTTQAA 565
 QY 536 NVITTNESITVLNNGNNINIKLDYRKELIAYNGWPGETDKNGKNGRLMLYKPTEDRTLL 595
 Db 566 TVTITGKDLA--TTGNN--NSIDSKEIAYNGWPGKDTTKNGRLMLYQPAEERLTL 623
 QY 596 SSGTNLKGDTQTKKLPFGSRPTPHAYNHLKRWSEMEGIPGELIWMNDHMINRTEKAE 655
 Db 624 SSGTNLKGDTQTKKLPFGSRPTPHAYNHLKRWSEMEGIPGELIWMNDHMINRTEKAE 683
 QY 656 NFOIKGSAVAVSRNYSIEGNWTVSNNAATFGVAVNQNONTICRSDMTGLTTCQKVDLT 715
 Db 684 NFOIKGSAVAVSRNYSIEGNWTVSNNAATFGVAVNQNONTICRSDMTGLTTCQKVDLT 743
 QY 716 DTKVINSIPCTOJNSINLTDNATANYKGLAKINGVTLNHSOFTLSNNAOTQGN--- 771
 Db 744 DDKVIASILEKTDISGVVDLADHAILNLTGLIATLNGNLSANEDTRYTVSHNATONGNLSLV 833
 QY 772 -----IRLSDKSTATVDNANLNGVHLTDS 796
 Db 804 GNAQATPKQATLNGNTSASGNASPNLSDBAVQNGSLISGNAKANVSHSALNGVSLADK 863
 QY 797 AQSFKNSHSHQIQDGITVILENATWMTSPDTLONTLNNSTITLINSAY----- 849
 Db 864 AVFHESSRFTQISGKOTALHKSSEWTLPSGTGLGNLNDATITLINSAYRHDAGA 923
 QY 850 -SASSNTPR-----RSLETTPTSAHRNTLTYNGKLSGQCTPOFSSLFYKSD 902
 Db 924 QTSALDAPRRRSRBSRSLSVTPPTSVSFRNTLTYNGKLSGQCTPOFSSLFYKSD 983
 QY 903 KIKLSNABGDYILSVRNTGKPELTLEQTLVESKONOPISDKLKFLENHVDAGAIRY 962
 Db 984 KIKLSNABGDYILSVRNTGKPELTLEQTLVESKONOPISDKLKFLENHVDAGAIRY 1043
 QY 963 KLVNDEFRILNPIKEOELHNDVRAE---QAER-----TLEAKOVEPTAKTQ-GE 1011
 Db 1044 QLIKDEFRILNPIKEOELHNDVRAE---QAER-----TLEAKOVEPTAKTQ-GE 1103
 QY 1012 PVTSSRAAPAAPPTLPOSLALAKAKAELTAFTQSKAKTKYRSKRAVPSDPLD 1071
 Db 1104 P-----ARQAGEVNV---GIMOAEKKEK-RVQADKOTALAKREAEKTRPATTFP-- 1149
 QY 1072 QSLFALMALEVIDAQQSBKORLAQEAQKOKODLISRYNSAISLSTVNSMLSV 1131
 Db 1150 -----FAKRRARDLPQLOPOPQOPQADLISRYNSAISLSTVNSMLSV 1195

QY 1132 QDELRLFPVDAQSAVINIADNRRTDSDAPRAVQOQKTLNRQIGYQKXLANGRIQAVF 1192
 Db 1196 QDELRLFPVDAQSAVINIADNRRTDSDAPRAVQOQKTLNRQIGYQKXLANGRIQAVF 1254
 QY 1192 SHSRSDNTPOQYKXNATLTMWSGPAQYQMGDLQFQVNVGTGISASQMAEQSRKXIHKA 1251
 Db 1255 SHSRSDNTPOQYKXNATLTMWSGPAQYQMGDLQFQVNVGTGISASQMAEQSRKXIHKA 1314
 QY 1252 INVGNASQYQFRIQLOGLQYFQVNRVYIERENYQSEEVRYKTPSLANRNNAGIRVDYT 1311
 Db 1315 LHYQIAPYRAGRGFGIEPHIGATRFVQKADRYENNVNATPGLAPNRYRAGIKADYS 1374
 QY 1312 FPPDNTISVKEYPFVNVYDVSNANVQTTNRLTVLQOPFRRTYQXEVGLKKEILHFOISAF 1371
 Db 1375 FKPAQHISITPLYSITDASGKVTREVTAVLAQDFKTRSAWGVNABIKGFTLSH 1434
 QY 1372 ISKSGSOLGQOQVGVKLYRW 1394
 Db 1435 AAAAGPQLEHQSAGTKLYRW 1457

Search completed: June 23, 2004, 17:33:45
 Job time : 72 secs

; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1394 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-645-655-2

Query Match 100.0%; Score 7273; DB 12; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKTYFRNLPLTACISLGIVSQAMAGHTYFIDYQYRDPFAENKGFVGAQNIKYNNKQ 60
 DB 1 MKKTYFRNLPLTACISLGIVSQAMAGHTYFIDYQYRDPFAENKGFVGAQNIKYNNKQ 60
 QY 61 GQIVGSMTKAPMIDSVSRNGVAALVENQYIVSAHVAVGYTVDPCAGANNPDQHRFT 120
 DB 61 GQIVGSMTKAPMIDSVSRNGVAALVENQYIVSAHVAVGYTVDPCAGANNPDQHRFT 120
 QY 121 YKIVKNNYKKNLHPEDDYNRPRLHKFVTEAPIDMTSNNGSTYSDETKYPERVRIG 180
 DB 121 YKIVKNNYKKNLHPEDDYNRPRLHKFVTEAPIDMTSNNGSTYSDETKYPERVRIG 180
 QY 181 SARQFWNDQDKGDQVAGAVHYLTAGTNHNGAGNGSYTLGGDVRAKAGEYGLPIAGSK 240
 DB 181 SARQFWNDQDKGDQVAGAVHYLTAGTNHNGAGNGSYTLGGDVRAKAGEYGLPIAGSK 240
 QY 241 GDSGSPMFTYAEKQKVLNGLIRGNPEPEKENGFLYRKSYFDELIRSDLRHTSLYTA 300
 DB 241 GDSGSPMFTYAEKQKVLNGLIRGNPEPEKENGFLYRKSYFDELIRSDLRHTSLYTA 300
 QY 301 GNGVYTSNDNGGOSTITQSGIPSEIKITLANMSLPEAKQKYNPRYDGPNIYSPRLN 360
 DB 301 GNGVYTSNDNGGOSTITQSGIPSEIKITLANMSLPEAKQKYNPRYDGPNIYSPRLN 360
 QY 361 NGBTIYFMDQKGSILFASDINGAGLYFEFGNTVPSNSCTWQAGIHYSENSVTYTK 420
 DB 361 NGBTIYFMDQKGSILFASDINGAGLYFEFGNTVPSNSCTWQAGIHYSENSVTYTK 420
 QY 421 VNGVEHRLSKTGKTLHVAKEKNGKSYVGDQKVLLEQADQGNQASBGLVYSR 480
 DB 421 VNGVEHRLSKTGKTLHVAKEKNGKSYVGDQKVLLEQADQGNQASBGLVYSR 480
 QY 481 GTVQMDKQPTOKETFEFGGRGLDINHSILTEKRIQNTDEGAMIVANNTQAANVTIT 540
 DB 481 GTVQMDKQPTOKETFEFGGRGLDINHSILTEKRIQNTDEGAMIVANNTQAANVTIT 540
 QY 541 GNSIIVPENGNNINKLDYRKEIAYNGMFEETDKKNGKGLIYPTTEEDTLISGGIN 600
 DB 541 GNSIIVPENGNNINKLDYRKEIAYNGMFEETDKKNGKGLIYPTTEEDTLISGGIN 600
 QY 601 LKGDITQTKGLFEFGGPTPHAYNHLNRKWSMGIPOGRIWDDHDMNRFKXENFOIK 660
 DB 601 LKGDITQTKGLFEFGGPTPHAYNHLNRKWSMGIPOGRIWDDHDMNRFKXENFOIK 660
 QY 661 GGSAYVSRVSSIEGNMTVSNANATFEVVPRQONTICTSDMTGLTTCOKRYDLTDTKYI 720
 DB 661 GGSAYVSRVSSIEGNMTVSNANATFEVVPRQONTICTSDMTGLTTCOKRYDLTDTKYI 720
 QY 721 NSIPKTIQNGSINLTNANAVKGLAKNGVNTLTNHSOCTLSNNAITQIGIRISDSNIA 780
 DB 721 NSIPKTIQNGSINLTNANAVKGLAKNGVNTLTNHSOCTLSNNAITQIGIRISDSNIA 780
 QY 781 TVNANINGNVHLTSAQPSLKNSHSHQIQDQKGTVTLENATWTMPDPTLONLTNN 840
 DB 781 TVNANINGNVHLTSAQPSLKNSHSHQIQDQKGTVTLENATWTMPDPTLONLTNN 840
 QY 841 STIINSAVSASANNTPRRSLETTPTSAHRENTLTVNGKLSGQGFQPTSLPGYK 900
 DB 841 STIINSAVSASANNTPRRSLETTPTSAHRENTLTVNGKLSGQGFQPTSLPGYK 900

QY 901 SDKILSNABGDYILSVRNGKPELTLEQLTVESKQNPLODKLFTLENDHVDAGAL 960
 DB 901 SDKILSNABGDYILSVRNGKPELTLEQLTVESKQNPLODKLFTLENDHVDAGAL 960
 QY 961 RYKLVKNDGFRFLNPIKEQELHNDLVRAEOAERTLEAKQVEPLANTQGEPEKVRRAA 1020
 DB 961 RYKLVKNDGFRFLNPIKEQELHNDLVRAEOAERTLEAKQVEPLANTQGEPEKVRRAA 1020
 QY 1021 PAAPFTLPPQSLNALAAKQAEITLAEQSKAKTKYRSRAVAFSDPLIDSLFLAEAA 1080
 DB 1021 PAAPFTLPPQSLNALAAKQAEITLAEQSKAKTKYRSRAVAFSDPLIDSLFLAEAA 1080
 QY 1081 LEVDAPOQSEKDRLAQEBEAKQKODLISRYNSALSELATVSMMLSVODELRLFLV 1140
 DB 1081 LEVDAPOQSEKDRLAQEBEAKQKODLISRYNSALSELATVSMMLSVODELRLFLV 1140
 QY 1141 DQASAVNTIADQKRRYDSAFRAYQOQKTMIRQIGVQKALANGRIGAVFSSHSINTF 1200
 DB 1141 DQASAVNTIADQKRRYDSAFRAYQOQKTMIRQIGVQKALANGRIGAVFSSHSINTF 1200
 QY 1201 DEOVNHAITLTWMSGFADQYQWGDLOFGVNVGTGISASMAEESRKIHRAAINYGMAISY 1260
 DB 1201 DEOVNHAITLTWMSGFADQYQWGDLOFGVNVGTGISASMAEESRKIHRAAINYGMAISY 1260
 QY 1261 QPRIGOLGIQPYGVVRYEIERENYQSEEVRYKTSIAFNYNAGIRVDYFTPTDNISV 1320
 DB 1261 QPRIGOLGIQPYGVVRYEIERENYQSEEVRYKTSIAFNYNAGIRVDYFTPTDNISV 1320
 QY 1321 KEFFVNVYDVSNANVQTTNLTVLQQPFGRITQWEVGLAEIILHFOISAFISKQSGQL 1380
 DB 1321 KEFFVNVYDVSNANVQTTNLTVLQQPFGRITQWEVGLAEIILHFOISAFISKQSGQL 1380
 QY 1381 GKQQWYKYGKYN 1394
 DB 1381 GKQQWYKYGKYN 1394

RESULT 3
 US-10-080-505-2
 ; Sequence 2, Application US/10080505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OR INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RF/DCE/DBR
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,731
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 1394
 ; TYPE: PRP
 ; ORGANISM: Haemophilus influenzae
 US-10-080-505-2

Query Match 100.0%; Score 7273; DB 14; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKTYFRNLPLTACISLGIVSQAMAGHTYFIDYQYRDPFAENKGFVGAQNIKYNNKQ 60
 DB 1 MKKTYFRNLPLTACISLGIVSQAMAGHTYFIDYQYRDPFAENKGFVGAQNIKYNNKQ 60
 QY 61 GQIVGSMTKAPMIDSVSRNGVAALVENQYIVSAHVAVGYTVDPCAGANNPDQHRFT 120
 DB 61 GQIVGSMTKAPMIDSVSRNGVAALVENQYIVSAHVAVGYTVDPCAGANNPDQHRFT 120
 QY 121 YKIVKNNYKKNLHPEDDYNRPRLHKFVTEAPIDMTSNNGSTYSDETKYPERVRIG 180
 DB 121 YKIVKNNYKKNLHPEDDYNRPRLHKFVTEAPIDMTSNNGSTYSDETKYPERVRIG 180

Db 121 YKIVKANNYKONLHPEEDYHNPRLHKFVTEAELDMTSNMNGSTYSDRITYEERVRIQ 180
 Qy 121 SGRQPMNDQKGDVAGAYHYLTAAGNTNMGAGNGSVYIGDVPKAEVGPPLIGSK 240
 Db 121 SGRQPMNDQKGDVAGAYHYLTAAGNTNMGAGNGSVYIGDVPKAEVGPPLIGSK 240
 Qy 241 GDSGPMFYDAEKOKMLINGILREGNPEEGENGPOLYVRSYFDSIPERDLHTSLYTRA 300
 Db 241 GDSGPMFYDAEKOKMLINGILREGNPEEGENGPOLYVRSYFDSIPERDLHTSLYTRA 300
 Qy 301 GNGVYTTISGNDGQGSITQKSGIPSEIKITLANNGLPKEXKRVNHPRYDGNITYSRLN 360
 Db 301 GNGVYTTISGNDGQGSITQKSGIPSEIKITLANNGLPKEXKRVNHPRYDGNITYSRLN 360
 Qy 361 NGELTYFMNDQKGSILFASDINOGAGLYFEENFTVSNNOQWQAGAHSENSYTWK 420
 Db 361 NGELTYFMNDQKGSILFASDINOGAGLYFEENFTVSNNOQWQAGAHSENSYTWK 420
 Qy 421 VNGVEHRLSKIGKGLTHVQAKGENKGSISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Db 421 VNGVEHRLSKIGKGLTHVQAKGENKGSISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Qy 481 GTVOLNDKQFDTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMTVNNTTOAMVTTT 540
 Db 481 GTVOLNDKQFDTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMTVNNTTOAMVTTT 540
 Qy 541 GNESTIVLPNNNNINKLDRKEIAYNCFGBTDKNNGNRNLTYKFTEDRLLSGTN 600
 Db 541 GNESTIVLPNNNNINKLDRKEIAYNCFGBTDKNNGNRNLTYKFTEDRLLSGTN 600
 Qy 601 LKGDITQTKGKLFEPSGRPTPHAVNHLNKSMESEIPOSEVWMDHMINRTEKAEFOIK 660
 Db 601 LKGDITQTKGKLFEPSGRPTPHAVNHLNKSMESEIPOSEVWMDHMINRTEKAEFOIK 660
 Qy 661 GGSAYSRNVSSIEGNTVSNANAFGVVPOQNTTCTRSBWTSLTTOQKVDLDTKVI 720
 Db 661 GGSAYSRNVSSIEGNTVSNANAFGVVPOQNTTCTRSBWTSLTTOQKVDLDTKVI 720
 Qy 721 NSIKRQJNGSINTLDMATANVKGLAKNGCNTLTJMSOFTLSNMTOLGNIRLSNDA 780
 Db 721 NSIKRQJNGSINTLDMATANVKGLAKNGCNTLTJMSOFTLSNMTOLGNIRLSNDA 780
 Qy 781 TVDNANLNGVHLTDSAQFSLKNSHFSHQJQDKGTVTLFNATWTPSDITLQNTLNN 840
 Db 781 TVDNANLNGVHLTDSAQFSLKNSHFSHQJQDKGTVTLFNATWTPSDITLQNTLNN 840
 Qy 841 STTILNSAIVASSNNTPRRRLTETETTPSAHRNNTLTNGKLSGQCFOTTSIPEYK 900
 Db 841 STTILNSAIVASSNNTPRRRLTETETTPSAHRNNTLTNGKLSGQCFOTTSIPEYK 900
 Qy 901 SDPKLNDAGEGYILSVANTGKEPETLEOLTJVESKONCPJSDKLTLENDHVDAGAL 960
 Db 901 SDPKLNDAGEGYILSVANTGKEPETLEOLTJVESKONCPJSDKLTLENDHVDAGAL 960
 Qy 961 RYKLVKNDGEFLNPIKEOELNDIVRABOERTLEAKOVEPTAKTJGEEKVSRPRA 1020
 Db 961 RYKLVKNDGEFLNPIKEOELNDIVRABOERTLEAKOVEPTAKTJGEEKVSRPRA 1020
 Qy 1021 RAAPFTTLPQOSILNLEKQMLTAELOKSAKTKYKRSRAVSPPLDOSLPALEAA 1080
 Db 1021 RAAPFTTLPQOSILNLEKQMLTAELOKSAKTKYKRSRAVSPPLDOSLPALEAA 1080
 Qy 1081 LEVIDAPOGSEKRLAOEAEKOKKODLSRNSNALSSEATVNSMLSVDEJDRLEFV 1140
 Db 1081 LEVIDAPOGSEKRLAOEAEKOKKODLSRNSNALSSEATVNSMLSVDEJDRLEFV 1140
 Qy 1141 DOAGSAWNTIAODKRRYDSAPRAYOQKTNLRQJGVQKLANGRIGAVFHSRSDNTF 1200
 Db 1141 DOAGSAWNTIAODKRRYDSAPRAYOQKTNLRQJGVQKLANGRIGAVFHSRSDNTF 1200
 Qy 1201 DEQVNEHRLSKIGKGLTHVQAKGENKGSISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Db 1201 DEQVNEHRLSKIGKGLTHVQAKGENKGSISVGDGKYLLEQADPOGNKAFSEIGVSGR 480

Qy 1261 OFRLQGLGIPYFGVARYFLERENYQSEEVKTPSLAFNRNAGIRVDYTPPTDNISV 1320
 Db 1261 OFRLQGLGIPYFGVARYFLERENYQSEEVKTPSLAFNRNAGIRVDYTPPTDNISV 1320
 Qy 1321 KPYFFNVYDVSNANQTTWNLTJVOQFGRVQKGVGKAEILHFOISAFLSKSGSOL 1380
 Db 1321 KPYFFNVYDVSNANQTTWNLTJVOQFGRVQKGVGKAEILHFOISAFLSKSGSOL 1380
 Qy 1381 GKCQWGVKLGKYM 1394
 Db 1381 GKCQWGVKLGKYM 1394

RESULT 4

US-10-080-505-7
 ; Sequence 7, Application US/10086505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RET/DCF/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 1395
 ; TYPE: PR1
 ; ORGANISM: Haemophilus influenzae
 US-10-080-505-7

Query Match 99.6%; Score 7246.5; DB 14; Length 1395;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1393; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

Qy 1 MKTIVRLNPLTACISIGVSOAMAGTYRGIDYQYEDPAENKKTGVGKONIKYVNO 60
 Db 1 MKTIVRLNPLTACISIGVSOAMAGTYRGIDYQYEDPAENKKTGVGKONIKYVNO 60
 Qy 61 GOLVGSATKAPMIDFVSVRNGVAAVYENQYIVSAVHNGYTDVDFGABGNNDQHRT 120
 Db 61 GOLVGSATKAPMIDFVSVRNGVAAVYENQYIVSAVHNGYTDVDFGABGNNDQHRT 120
 Qy 121 YKIVKANNYKONLHPEEDYHNPRLHKFVTEAELDMTSNMNGSTYSDRITYEERVRIQ 180
 Db 121 YKIVKANNYKONLHPEEDYHNPRLHKFVTEAELDMTSNMNGSTYSDRITYEERVRIQ 180
 Qy 181 SGRQPMNDQKGDVAGAYHYLTAAGNTNMGAGNGSVYIGDVPKAEVGPPLIGSK 240
 Db 181 SGRQPMNDQKGDVAGAYHYLTAAGNTNMGAGNGSVYIGDVPKAEVGPPLIGSK 240
 Qy 241 GDSGPMFYDAEKOKMLINGILREGNPEEGENGPOLYVRSYFDSIPERDLHTSLYTRA 300
 Db 241 GDSGPMFYDAEKOKMLINGILREGNPEEGENGPOLYVRSYFDSIPERDLHTSLYTRA 300
 Qy 301 GNGVYTTISGNDGQGSITQKSGIPSEIKITLANNGLPKEXKRVNHPRYDGNITYSRLN 360
 Db 301 GNGVYTTISGNDGQGSITQKSGIPSEIKITLANNGLPKEXKRVNHPRYDGNITYSRLN 360
 Qy 361 NGELTYFMNDQKGSILFASDINOGAGLYFEENFTVSNNOQWQAGAHSENSYTWK 420
 Db 361 NGELTYFMNDQKGSILFASDINOGAGLYFEENFTVSNNOQWQAGAHSENSYTWK 420
 Qy 421 VNGVEHRLSKIGKGLTHVQAKGENKGSISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Db 421 VNGVEHRLSKIGKGLTHVQAKGENKGSISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Qy 481 GTVOLNDKQFDTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMTVNNTTOAMVTTT 540

```

Db      481 GTVQNDKQPDIDKFFPGRGRLDLSLFFKQCTGEGMAYNHATTQANVIT 540
Qy      541 GNEISIVLPGNNINKLDYRKELIANGNGEGEDKNGKNGRLNIYKPTEDRELLISGGTN 600
Db      541 GNEISIVLPGNNINKLDYRKELIANGNGEGEDKNGKNGRLNIYKPTEDRELLISGGTN 600
Qy      601 JGDDITVTKGKLFPSGRPTPHAVNHLNKRSGEMEGIPQGEIWMHDNINTEFAENFOIK 660
Db      601 JGDDITVTKGKLFPSGRPTPHAVNHLNKRSGEMEGIPQGEIWMHDNINTEFAENFOIK 660
Qy      661 GSAVAVSRNVSIEGMMTVSNANATGEGVPOONTICHSMTGLTCCXKDLDTDTVI 720
Db      661 GSAVAVSRNVSIEGMMTVSNANATGEGVPOONTICHSMTGLTCCXKDLDTDTVI 720
Qy      721 NSIPKTOGNSINLTNATANYKGLAKNGVTLTHSQFTLSNNATQIGNIRLSNSTA 780
Db      721 NSIPKTOGNSINLTNATANYKGLAKNGVTLTHSQFTLSNNATQIGNIRLSNSTA 780
Qy      781 TVDNANLNGNVHLTDSAPSLNKHSHSHQIQGKTTVLENATVWPSDTTQNTLNN 840
Db      781 TVDNANLNGNVHLTDSAPSLNKHSHSHQIQGKTTVLENATVWPSDTTQNTLNN 840
Qy      841 STITNSAVSASSNNPT--RBSLETTTCTSAHRENTLVNKGSGGTFQPTSSJG 898
Db      841 STITNSAVSASSNNPT--RBSLETTTCTSAHRENTLVNKGSGGTFQPTSSJG 898
Qy      899 YKSDKLKLSNDAEGDYLISVNTKEPELTLEQLLVESKDQPLSDKLKTLNHDVAG 958
Db      901 YKSDKLKLSNDAEGDYLISVNTKEPELTLEQLLVESKDQPLSDKLKTLNHDVAG 960
Qy      959 AIRYKLVNDSGERLNLPIKEOEHLNDVPAQDAEFTLENAQVEPTAKCTGEPKYSRR 1018
Db      961 AIRYKLVNDSGERLNLPIKEOEHLNDVPAQDAEFTLENAQVEPTAKCTGEPKYSRR 1020
Qy      1019 AARAAFPDTPDQSLNALBAKQALTAPOKSKATKVKCSKAVSDPLDQSLPALE 1078
Db      1021 AARAAFPDTPDQSLNALBAKQALTAPOKSKATKVKCSKAVSDPLDQSLPALE 1080
Qy      1079 AALEVITAPQOSEKRLAQEBAKQKQKOLISRYSNALSLSATVNSMLSVODELRL 1138
Db      1081 AALEVITAPQOSEKRLAQEBAKQKQKOLISRYSNALSLSATVNSMLSVODELRL 1140
Qy      1139 FVDOQASVMTNIAQDKRYDSDAARAYQOQKTNLFQIYQKALANRIGAVSHSRBDN 1198
Db      1141 FVDOQASVMTNIAQDKRYDSDAARAYQOQKTNLFQIYQKALANRIGAVSHSRBDN 1199
Qy      1199 TDEOVKNHATLTMSGFAQYOMGDLFGVNVGTGISASKVAEBSKIRKAIYGVNA 1258
Db      1200 TDEOVKNHATLTMSGFAQYOMGDLFGVNVGTGISASKVAEBSKIRKAIYGVNA 1259
Qy      1259 SYQFRLGQIGIOPYFGVNYPIERENTYCSSEVAVKTPSLAENRYNAGIRVDTFTPTNI 1318
Db      1260 SYQFRLGQIGIOPYFGVNYPIERENTYCSSEVAVKTPSLAENRYNAGIRVDTFTPTNI 1319
Qy      1319 SVKPYFVNYVDVSNANVQITVNLVQOPRGYKQKVGKALTEHPQISAFISKSQS 1378
Db      1320 SVKPYFVNYVDVSNANVQITVNLVQOPRGYKQKVGKALTEHPQISAFISKSQS 1379
Qy      1379 QLGKQONVGVLAGYRW 1394
Db      1380 QLGKQONVGVLAGYRW 1395

RESULT 5
; Sequence 9, Application US/10080505
; Publication No. US20030073166A1
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505

; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,396
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (568)..(568)
; OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (1702)..(1702)
; OTHER INFORMATION: "n" at position 1702 can be any base.
US-10-080-505-9

Query Match      84.1%; Score 6115; DB 14; Length 1434;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 10;

Qy      1 MKKTVRLNPLFACSGIIVSOAMAGHTYFGIDYQYRDFRAEKKFTVGAQNTKVTNK 60
Db      1 MKKTVRLNPLFACSGIIVSOAMAGHTYFGIDYQYRDFRAEKKFTVGAQNTKVK 60
Qy      61 GQLVGTSMTKAMIDFVSVSRNGVAALVENQYIVSAVANGYTDVDPGAENNDQHRFT 120
Db      61 GQLVGTSMTKAMIDFVSVSRNGVAALVGDYIVSAVANGYTNVDPFGAENNDQHRFT 120
Qy      61 GQLVGTSMTKAMIDFVSVSRNGVAALVGDYIVSAVANGYTNVDPFGAENNDQHRFT 120
Db      121 YKIKVNNYKKNLHPYEDYHNPLAKVTIAPIDMTSNNGSYSDRTYPERVETG 180
Qy      121 YKIKVNNYKKNLHPYEDYHNPLAKVTIAPIDMTSNNGSYSDRTYPERVETG 180
Db      121 YKIKVNNYKKNLHPYEDYHNPLAKVTIAPIDMTSNNGSYSDRTYPERVETG 180
Qy      181 SGRQFMNDQDQVAGAYHLTAGTNRGAGNGSYLAGDVAKGEXGPLPIASGK 240
Db      181 SGRQFMNDQDQVAGAYHLTAGTNRGAGNGSYLAGDVAKGEXGPLPIASGK 240
Qy      241 GDGSSPWFYDAEKQKMLINGI..REGNPEKENGFOYVKSYPDEIIFERDLHSLYTRA 300
Db      241 GDGSSPWFYDAEKQKMLINGI..REGNPEKENGFOYVKSYPDEIIFERDLHSLYTRA 300
Qy      301 GNGYTTISGNDKQGSITQKSGTPESEIKITLANSPLPKKEXYHNPREDGNTIYSPRLN 360
Db      301 GNGYTTISGNDKQGSITQKSGTPESEIKITLANSPLPKKEXYHNPREDGNTIYSPRLN 360
Qy      361 NGETLYEMQKQSLIFASDINOGAGLYFEGNFTVSPNSQTMQAGIHWSENSVTYWK 420
Db      361 NGETLYEMQKQSLIFASDINOGAGLYFEGNFTVSPNSQTMQAGIHWSENSVTYWK 420
Qy      421 VNGEYDRSLKIGKGTLLHVOAKENGKISVGDQVILBQACADQGNKQAFSEIGLVSR 480
Db      421 VNGEYDRSLKIGKGTLLHVOAKENGKISVGDQVILBQACADQGNKQAFSEIGLVSR 480
Qy      481 GTVQNDKQPDIDKFFPGRGRLDLSLFFKQCTGEGMAYNHATTQANVIT 540
Db      481 GTVQNDKQPDIDKFFPGRGRLDLSLFFKQCTGEGMAYNHATTQANVIT 540
Qy      541 GNEISIVLPGNNINKLDYRKELIANGNGEGEDKNGKNGRLNIYKPTEDRELLISGGTN 600
Db      541 GNEISIVLPGNNINKLDYRKELIANGNGEGEDKNGKNGRLNIYKPTEDRELLISGGTN 600
Qy      600 NKAGDITVTKGKLFPSGRPTPHAVNHLNKRSGEMEGIPQGEIWMHDNINTEFAENFOIK 659
Db      600 NKAGDITVTKGKLFPSGRPTPHAVNHLNKRSGEMEGIPQGEIWMHDNINTEFAENFOIK 658
Qy      659 KGSAAVSRNVSIEGMMTVSNANATGEGVPOONTICHSMTGLTCCXKDLDTDTVI 719
Db      659 KGSAAVSRNVSIEGMMTVSNANATGEGVPOONTICHSMTGLTCCXKDLDTDTVI 718
Qy      720 NSIPKTOGNSINLTNATANYKGLAKNGVTLTHSQFTLSNNATQIGNIRLS---- 775

```

```

Db 715 ILSITTOIMSGINLTDNAIYINGLAKNGAVTLINSHQETLSNNAITGINKISNEAN 778
Qy 776 ---DNST-----ATVNAINAMNVAHTLSAQFS 800
Db 779 AAVNNATIMDYNALADTSFLSNQATIGTISLHQQAIVDANAMNVAHTLSAQFS 838
Qy 801 LKXSHFSHQIOGKGTIVLENATWTPSDTTLQNTLNNSTITLNSAYSASSANTP-RR 859
Db 839 LKXSHFSHQIOGKGTIVLENATWTPSDTTLQNTLNNSTITLNSAYSASSANTP-RR 898
Qy 860 RSLFETTPSAHRRNTLVNKGKSGGTGPTSLFGYSDKLKLSNDAEDDTLSVR 919
Db 899 RSLFETTPSAHRRNTLVNKGKSGGTGPTSLFGYSDKLKLSNDAEDDTLSVR 958
Qy 920 NTGKEPETLEQLTVESKNOPLSDKFTLENDHYDAGALRYLVNKGDEPRLANPIKE 979
Db 959 NTGKEPETLEQLTVESKNOPLSDKFTLENDHYDAGALRYLVNKGDEPRLANPIKE 1018
Qy 980 QELANDLVRABOAKERTLEAKOYEPTAKTGTGEPKYSRRAPAAAFEDTLFDOSLINALEA 1039
Db 1019 QELANDLVRABOAKERTLEAKOYEPTAKTGTGEPKYSRRAPAAAFEDTLFDOSLINALEA 1075
Qy 1040 KQALTAETOKSKAKTKKYSRGRAY---FSDPLDQS-LFALPALEVIDAPOSERKDL 1095
Db 1076 KQALTAETOKSKAKTKKYSRGRAY---FSDPLDQS-LFALPALEVIDAPOSERKDL 1135
Qy 1096 AQ-EEAEKOROKOLISRYNSALSELATVNSMKLWODELDELFDVDAQSAVMTNIAOD 1154
Db 1136 AQ-EEAEKOROKOLISRYNSALSELATVNSMKLWODELDELFDVDAQSAVMTNIAOD 1195
Qy 1155 KRRYSDAFRAYOQOKTNRLOIGVYKALANGRIKAVFSHSDSDTTPBOYKNAHTLTKMS 1214
Db 1196 KRRYSDAFRAYOQOKTNRLOIGVYKALANGRIKAVFSHSDSDTTPBOYKNAHTLTKMS 1254
Qy 1215 GFAOYOMGDLQFVNVGTCISASKAEHOSRXIHRKAIYVGNASVYOGLQGLOPYFG 1274
Db 1255 GFAOYOMGDLQFVNVGTCISASKAEHOSRXIHRKAIYVGNASVYOGLQGLOPYFG 1314
Qy 1275 VRRYIERENTOSEEVKTPSLAFNRNAGIRVDYTFPTDNISSKPYFVNVYDVNSA 1334
Db 1315 VRRYIERENTOSEEVKTPSLAFNRNAGIRVDYTFPTDNISSKPYFVNVYDVNSA 1374
Qy 1335 NVQTYNLTWIOQFPGYKQVGIKATLHPQISAFISKSGSGLKQONVYKGRW 1394
Db 1375 NVQTYNLTWIOQFPGYKQVGIKATLHPQISAFISKSGSGLKQONVYKGRW 1434

```

RESULT 6

```

US-10-680-505-17
; Sequence 17, Application US/10080505
; Publication No. US20030073166A1
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph M.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RET/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-680-505-17

```

Query Match 80.1%; Score 5827.5; DB 14; Length 1411;
 Best Local Similarity 79.5%; Pred. No. 0;
 Matches 1133; Conservative 107; Mismatches 139; Indels 47; Gaps 12;

```

Qy 1 MKKTVRLNPLTACISLGIVSQAMAGHTYFGIDYQYRDFAEENKGFVGAQNTKYTNKQ 60
Db 1 MKKTVRLNPLTACISLGIVSQAMAGHTYFGIDYQYRDFAEENKGFVGAQNTKYTNKQ 60
Qy 61 GQIVGSMKAKAMIDFVSRRNGVAAVENVYIVSVANVGYTVDVFAEENNDQHRFT 120
Db 61 GQIVGSMKAKAMIDFVSRRNGVAAVENVYIVSVANVGYTVDVFAEENNDQHRFT 120
Qy 121 YKIVKANNKKNIDHPYEDDYHNPRLHKFVTEAPADMTSNMNGSTYSDRTKYBERVIG 180
Db 121 YKIVKANNKKNIDHPYEDDYHNPRLHKFVTEAPADMTSNMNGSTYSDRTKYBERVIG 180
Qy 181 SGQVPRNDQKQVQVAGAVHYTAGTNHNRGAGNYSTLGSDYRKAAGEYGLPIAGSK 240
Db 181 SGQVPRNDQKQVQVAGAVHYTAGTNHNRGAGNYSTLGSDYRKAAGEYGLPIAGSK 240
Qy 241 GDSGSPMFLYAEKOKLNLINGLRBGNPEBKENGPOLYKSPYEFEDLHLSLYTRA 300
Db 241 GDSGSPMFLYAEKOKLNLINGLRBGNPEBKENGPOLYKSPYEFEDLHLSLYTRA 300
Qy 301 GNGVYITSGNDNGQSITOKSGIPSEKITTANSLPLAKKCKVHNPRYDGNPTYSBRN 360
Db 301 GNGVYITSGNDNGQSITOKSGIPSEKITTANSLPLAKKCKVHNPRYDGNPTYSBRN 360
Qy 361 NGELTYFMDQKGLIFASDINGAGLYFEGNTFTVPSNQNQWQAGIHSNSTVTWK 420
Db 361 NGELTYFMDQKGLIFASDINGAGLYFEGNTFTVPSNQNQWQAGIHSNSTVTWK 420
Qy 421 VNGVEHRLSKIGKTLHVOAKENKGSISVQDKVTLLEQOADDQNNQARSEIGLVGR 480
Db 421 VNGVEHRLSKIGKTLHVOAKENKGSISVQDKVTLLEQOADDQNNQARSEIGLVGR 480
Qy 481 GTVQNDNDKQPTDKFFYFGRGRGLDNLNHSITFKRIQNTDSBAMIVNNTQAAVTLT 540
Db 481 GTVQNDNDKQPTDKFFYFGRGRGLDNLNHSITFKRIQNTDSBAMIVNNTQAAVTLT 540
Qy 541 GNSBYLPRGNKINLKYRKELIYNGWPEGDDKNNKNGRNLTKYPTEDDTLLSGTN 600
Db 541 GNSBYLPRGNKINLKYRKELIYNGWPEGDDKNNKNGRNLTKYPTEDDTLLSGTN 600
Qy 601 LKQDITQYTKLFFSGRPTPHAVNHLNKSWESEMGIPQGIIVMDHDMINRFKAENFOIK 660
Db 601 LKQDITQYTKLFFSGRPTPHAVNHLNKSWESEMGIPQGIIVMDHDMINRFKAENFOIK 660
Qy 661 GGSIVSRVSSIBGNWTVSNANATFVYPRQONTICTSDMTGLTTCQKXVDTTKVI 720
Db 661 GGSIVSRVSSIBGNWTVSNANATFVYPRQONTICTSDMTGLTTCQKXVDTTKVI 720
Qy 721 NSIPKQINGSLNLDNATLVNKGKAKNGAVTLTNHSDFTLSNNAITGINKISNEAN 780
Db 721 NSIPKQINGSLNLDNATLVNKGKAKNGAVTLTNHSDFTLSNNAITGINKISNEAN 780
Qy 781 TVDNAMINAMNVAHTLSAQFSIKNSHFSHQIOGKGTIVLENATWTPSDTTLQNTLNN 840
Db 781 TVDNAMINAMNVAHTLSAQFSIKNSHFSHQIOGKGTIVLENATWTPSDTTLQNTLNN 840
Qy 841 STITANSAYSASSANTP-RRSLETTTPSAHRRNTLVNKGKSGGTGPTSLFGY 899
Db 841 STITANSAYSASSANTP-RRSLETTTPSAHRRNTLVNKGKSGGTGPTSLFGY 899
Qy 899 KSDKLKLSNDAEDDTLSVRNTGKEPETLEQLTVESKNOPLSDKFTLENDHYDAGAL 959
Db 899 KSDKLKLSNDAEDDTLSVRNTGKEPETLEQLTVESKNOPLSDKFTLENDHYDAGAL 959
Qy 959 LRYLVNKGDEPRLANPIKEQELANDLVRABOAKERTLEAKOYEPTAKTGTGEPKYSRRA 1019
Db 959 LRYLVNKGDEPRLANPIKEQELANDLVRABOAKERTLEAKOYEPTAKTGTGEPKYSRRA 1019
Qy 1019 ARAAFPTLFDOSLINALEAKQALTEKOKSKA-----KTKVSKR---AV 1064
Db 1019 ARAAFPTLFDOSLINALEAKQALTEKOKSKA-----KTKVSKR---AV 1064

```

QY 1065 FSDPILDDSL-----FA-----LEALEVIDADPOSEKORLAQSEAKKOROKD 1108
 Db 1070 LFAIYKETAQIDFANAKLAQNLITQOLEKALAV---AQQEKEKRAQOQAROKOKD 1126
 QY 1109 LISRYNSALSELATVNSMLSVODELRLFVDOASAVMTIADQKRRYDAPFAVQO 1168
 Db 1127 LISRSNSALSELATVNSMLSVODELRLFVDOASAVMTIADQKRRYDAPFAVQO 1185
 QY 1169 QKTNIROIQVQKALANGRIGAVSHSRSDNTPEBOVKNEALITMMSGPAQYQMGDIQFQV 1228
 Db 1186 QKTNIROIQVQKALANGRIGAVSHSRSDNTPEBOVKNEALITMMSGPAQYQMGDIQFQV 1245
 QY 1229 NYGTGISAKMAEESQSKIRKAIYGVNAYSQFRLGOLGCPYGRNRFETIRENYSQSE 1288
 Db 1246 NYGTGISAKMAEESQSKIRKAIYGVNAYSQFRLGOLGCPYGRNRFETIRENYSQSE 1305
 QY 1289 ERYVKTPELAFNRNAGIRVDYFTPTDNISVXPYFFVYVYVDSANAVQTTVALMYLQOP 1348
 Db 1306 ERYVKTPELAFNRNAGIRVDYFTPTDNISVXPYFFVYVYVDSANAVQTTVALMYLQOP 1365
 QY 1349 EGRYQKEVGLKAEILHFOISAFISKSQSQOLGQONVGVKIGTRM 1394
 Db 1366 EGRYQKEVGLKAEILHFOISAFISKSQSQOLGQONVGVKIGTRM 1411

RESULT 7
 US-10-080-505-13
 ; Sequence 13, Application US/10080505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph M.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RET/DCR/DBR
 ; CURRENT APPLICATION NUMBER: US/10/080, 505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296, 791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839, 996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patentia version: 3.1
 ; SEQ ID NO 13
 ; LENGTH: 1436
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-10-080-505-13

Query Match 79.8%; Score 5807; DS 14; Length 1436;
 Best Local Similarity 79.3%; Pred. No. 0;
 Matches 1145; Conservative 78; Mismatches 154; Indels 68; Gaps 14;
 QY 1 MKKTVFLNLTACISLIGVSAQAAGHTYFSDIYYRDPANNGKFTTGAQNIKYKQ 60
 Db 1 MKKTVFLNLTACISLIGVSAQAAGHTYFSDIYYRDPANNGKFTTGAQNIKYKQ 60
 QY 61 GGLVGTSTKAPMIDSVSRNGVALVEXQYVSAHVAVGYTVDGAEQNNPQDHF 120
 Db 61 GGLVGTSTKAPMIDSVSRNGVALVEXQYVSAHVAVGYTVDGAEQNNPQDHF 120
 QY 121 YKIVGRNNYKKDLHFEEDYHNPFLKFTVTRAPIDMTSNNGSTYSDETKYPERVIG 180
 Db 121 YKIVGRNNYKKDLHFEEDYHNPFLKFTVTRAPIDMTSNNGSTYSDETKYPERVIG 180
 QY 181 SGROWRNDQD-----KQCVAGAHYITNGNTHNORAGNGYISLGQDYKRAAGYGLP 235
 Db 181 SGROWRNDQD-----KQCVAGAHYITNGNTHNORAGNGYISLGQDYKRAAGYGLP 240
 QY 236 IAGSGDGGSPFPIYAAKQKYLINGILREGNPFEGKNGPOLVRSYF--DEIFPRDLHT 294
 Db 241 TGVSTGDSGSPFPIYAAKQKYLINGILREGNPFEGKNGPOLVRSYF--DEIFPRDLHT 300
 QY 295 SLYTBAGNGVYITSGNDGQSGITQ---KSGISEIKITLANMSLPLKEXDKVHPRY-- 349
 Db 295 SLYTBAGNGVYITSGNDGQSGITQ---KSGISEIKITLANMSLPLKEXDKVHPRY-- 349

Db 301 TELPRSNHNGHFTSNNGNGITVQTQINKEVSNPQKRTVQLFNEALKKEX--EPVYAA 358
 QY 350 DGNPIYSPLINNGETLYPMQKQOSLIPASDINGAGGLYFEGNFTVSNNGMQTQAGI 409
 Db 359 GGVNAYEPLINNGKTIYCGDNGTGLIENINQAGAGLYFEQNGFTVSNNAITQAGV 418
 QY 410 HYSNNTVTKVNGVYEHRLSKIGKTLHVOAKKNGNGSISVQDGVYILEQADQSKQ 469
 Db 419 HVGSDSTVTKVNGVYEHRLSKIGKTLHVOAKKNGNGSISVQDGVYILEQADQSKQ 478
 QY 470 AFSHIGLVSGRGTVQNDXOPTDKYFGFRGSLDINGSLTPFR--QNTDEGAMVNH 529
 Db 479 AFKXGVIGSRATVQJNSADQVDPNNITFGFRGSLDINGSLTPFR--QNTDEGAMVNH 538
 QY 530 NTPQANVTITGNESIVLP--NGNNIKLDYRKEIAVNGFEGTDKRNKNGRLNLT--YKPT 568
 Db 539 NTPQANVTITGNESITAPSKNNIKLDYSKEIAVNGFEGTDKRNKNGRLNLT--YKPT 598
 QY 589 EDRITLLSGTILKXKDLITQIKGLFESGRPTPHANHLNKRMEGEPDQGIYWDHMI 648
 Db 599 EDRITLLSGTILKXKDLITQIKGLFESGRPTPHANHLN--RPNEL--GRQGVAVTDDMI 656
 QY 649 NRTFKANFQIKGSAVSRVSSIEGNTVSNNAATFGVYVNOQNTICTSDMTGLTT 708
 Db 657 TRTFKAMFQIKGSAVSRVSSIEGNTVSNNAATFGVYVNOQNTICTSDMTGLTT 726
 QY 709 CQKVDLTDKTVNSIPKQJQNGSINTDVAIVNKGAKLAKNGVTLTNHSQTTLSNNAQ 768
 Db 717 CKYVDLTDKTVNSIPKQJQNGSINTDVAIVNKGAKLAKNGVTLTNHSQTTLSNNAQ 776
 QY 769 IGRIRLSDNSTVNNALNGNHLVDSQPSLKNSHSHQIGDGTVTLENAVTMP 828
 Db 777 TGNQJLSNNAATVNNALNGNHLVDSQPSLKNSHSHQIGDGTVTLENAVTMP 836
 QY 829 SDITLQNLTNASTITLNSAVSASNNRTR--RSLTETPTPTSAERFNTLVNKGSLGQ 887
 Db 837 SDATLQNLTNASTITLNSAVSASNNAPRHRSLTETPTPTSAERFNTLVNKGSLGQ 896
 QY 888 GTFOFTSSLPFGYSKDKLSDNABGYIISVNTNGKEPTELEUTLVYSKQNPISDKL 947
 Db 897 GTFOFTSSLPFGYSKDKLSDNABGYIISVNTNGKEPTELEUTLVYSKQNPISDKL 956
 QY 948 FTLENDVHAGALRYKLVNDGSEFRLNPIKXQELHNDIVARQABRTLEAKQVEPTAKT 1007
 Db 957 FTLENDVHAGALRYKLVNDGSEFRLNPIKXQELHNDIVARQABRTLEAKQVEPTAKT 1016
 QY 1008 QTEPKVRSRARAAPEDTLPPQSLNALAEKQKELTAETQSKAKTKVNSKRAVSD 1067
 Db 1017 QTSNARVRSK--RAVFDTLPPQSLDVLQAEQVEPTAEKQKKA--KVRSKRAVSD 1071
 QY 1068 PILDQS-----LFALEALEVID 1085
 Db 1072 TLPPQSLDVLQAEQVEPTAEKQKAKVRSKRAAREFDTLPSRLKVLNLEVLN 1131
 QY 1086 APQOSEKRLAQEAEKOKOKDLISYNSALSELATVNSMLSVODELRLFVDOAS 1145
 Db 1132 AQOQYKKE---PQDEKOKOKDLISYNSALSELATVNSMLSVODELRLFVDOAS 1188
 QY 1146 AWTNIAQDKRRYDSDAFRAYQOKTNIROIQVQKALANGRIGAVSHSRSDNTPEBOVK 1205
 Db 1189 AWTNIAQDKRRYDSDAFRAY--QOKTNIROIQVQKALANGRIGAVSHSRSDNTPEBOVK 1247
 QY 1206 NHATLTMSGPAQYQMGDIQFQVNGGTGISAKMAEESQSKIRKAIYGVNAYSQFRLG 1265
 Db 1248 NHATLTMSGPAQYQMGDIQFQVNGGTGISAKMAEESQSKIRKAIYGVNAYSQFRLG 1307
 QY 1266 QLGIOYFPGVNRYPFERENYQSEEVVKTPELAFNRNAGIRVDYFTPTDNISVXPYFF 1325
 Db 1308 QLGIOYFPGVNRYPFERENYQSEEVVKTPELAFNRNAGIRVDYFTPTDNISVXPYFF 1367
 QY 1366 VNYVDSANAVQTTVALMYLQOPFRYQKEVGLKAEILHFOISAFISKSQSQOLGQON 1385
 Db 1368 VNYVDSANAVQTTVALMYLQOPFRYQKEVGLKAEILHFOISAFISKSQSQOLGQON 1427

QY 1386 VGVKLGXRM 1394
 Db 1428 VGVKLGXRM 1436

RESULT 8

US-10-080-505-15
 ; Sequence 15, Application US/10080505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RT/DCG/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080.505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 1391
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-10-080-505-15

Query Match 72.9%; Score 5303.5; DB 14; Length 1391;

Best Local Similarity 74.8%; Pred. No. 0;
 Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

QY 1 MKTVEINFLTACISLGIYSQAMAGHTYEGIDYCYHREKRGKFTYGAQNTKYNKQ 60
 Db 1 MKTVEINFLTACISLGIYSQAMAGHTYEGIDYCYHREKRGKFTYGAQNTKYNKQ 60
 QY 61 GQVGTSMKAPMIDFSVSRNGVALVENQYIVSVANNQYTDVFGAAGNPDQRF 120
 Db 61 GEMGTMMKGVPMPLSSMVGAGYSTLISEQHLISVANNGYQVDFGAGNPDQRF 120
 QY 121 YKTYKNNYKCNLHPEDDYHNPRLHKVTYAPIDMTSNMGSYSTDRITKYERYRIG 180
 Db 121 YKTYKNNYKCNLHPEDDYHNPRLHKVTYAPIDMTSNMGSYSTDRITKYERYRIG 180
 QY 121 YKTYKNNYKCNLHPEDDYHNPRLHKVTYAPIDMTSNMGSYSTDRITKYERYRIG 180
 Db 121 YKTYKNNYKCNLHPEDDYHNPRLHKVTYAPIDMTSNMGSYSTDRITKYERYRIG 180
 QY 181 SGQFWEINDGX--SDQVAGAYHLLTANNTNORAGNGSYLGEYKRAQBYPLTIA 238
 Db 181 SGQFWEINDGX--SDQVAGAYHLLTANNTNORAGNGSYLGEYKRAQBYPLTIA 238
 QY 179 SGQFWEINDGX--SDQVAGAYHLLTANNTNORAGNGSYLGEYKRAQBYPLTIA 238
 Db 179 SGQFWEINDGX--SDQVAGAYHLLTANNTNORAGNGSYLGEYKRAQBYPLTIA 238
 QY 239 SKGDSGPMFIYDAEKOKMLNGLRENGPFGKENGFOYKSYDELFFREDLHSLYT 298
 Db 239 SKGDSGPMFIYDAEKOKMLNGLRENGPFGKENGFOYKSYDELFFREDLHSLYT 298
 QY 238 SFGDSGPMFIYDAEKOKMLNGLRENGPFGKENGFOYKSYDELFFREDLHSLYT 297
 Db 238 SFGDSGPMFIYDAEKOKMLNGLRENGPFGKENGFOYKSYDELFFREDLHSLYT 297
 QY 239 RAGNGVYITSGDNG----QGSITQKSGPBEIKITLANSI-PIKEKDYNNPRYDQ 352
 Db 239 RAGNGVYITSGDNG----QGSITQKSGPBEIKITLANSI-PIKEKDYNNPRYDQ 352
 QY 238 TNAEYRENIQSDHNRVATIKSTLPKKAIOPRIVGLYDMSQLDAREKDESPSYKGP 357
 Db 238 TNAEYRENIQSDHNRVATIKSTLPKKAIOPRIVGLYDMSQLDAREKDESPSYKGP 357
 QY 353 NTVSPRLNGETLYMDQKOSLIFASDINOAGSLFPEENFVSPNSQ--TMQAGSIHY 411
 Db 353 NTVSPRLNGETLYMDQKOSLIFASDINOAGSLFPEENFVSPNSQ--TMQAGSIHY 411
 QY 358 NTVSPRLNGETLYMDQKOSLIFASDINOAGSLFPEENFVSPNSQ--TMQAGSIHY 417
 Db 358 NTVSPRLNGETLYMDQKOSLIFASDINOAGSLFPEENFVSPNSQ--TMQAGSIHY 417
 QY 412 SENSTVTKWNGVEHRLSKIGKTLFVQANKENGKSIYSGDKVILEQADQGNKQAF 471
 Db 412 SENSTVTKWNGVEHRLSKIGKTLFVQANKENGKSIYSGDKVILEQADQGNKQAF 471
 QY 418 GEESTVEMQVHNPEEDRLSKIGKTLFVQANKENGKSIYSGDKVILEQADQGNKQAF 477
 Db 418 GEESTVEMQVHNPEEDRLSKIGKTLFVQANKENGKSIYSGDKVILEQADQGNKQAF 477
 QY 472 SEIGVSGGVQLNDQKQPTDKYFGRGRGLDNGHSLTKYQNTDEGANTVNAHT 531
 Db 472 SEIGVSGGVQLNDQKQPTDKYFGRGRGLDNGHSLTKYQNTDEGANTVNAHT 531
 QY 478 KEVIVSGGVQLNDQKQPTDKYFGRGRGLDNGHSLTKYQNTDEGANTVNAHT 537
 Db 478 KEVIVSGGVQLNDQKQPTDKYFGRGRGLDNGHSLTKYQNTDEGANTVNAHT 537
 QY 532 TQANVTITGNEISYLPNGANNKLDYRKELIAYNGMGSEDDKKNKNGRLNLYFETEDR 591
 Db 532 TQANVTITGNEISYLPNGANNKLDYRKELIAYNGMGSEDDKKNKNGRLNLYFETEDR 591
 QY 538 SQANVTITGNEISYLPNGANNKLDYRKELIAYNGMGSEDDKKNKNGRLNLYFETEDR 593
 Db 538 SQANVTITGNEISYLPNGANNKLDYRKELIAYNGMGSEDDKKNKNGRLNLYFETEDR 593
 QY 592 TLLSGTINLKGDLITQTKGLPFSGRPTPHAYNHLNRKSEKGIPOGEIYVMDHDMINT 651
 Db 592 TLLSGTINLKGDLITQTKGLPFSGRPTPHAYNHLNRKSEKGIPOGEIYVMDHDMINT 651

Db 594 HLLSGTINLKGDLITQTKGLPFSGRPTPHAYNHLNRKSEKGIPOGEIYVMDHDMINT 653
 QY 652 FAENFOIKGSAVAVSRVSSIEGNWTVSNANAFGVVPMQNTICTRSDMTGLTTCOK 711
 Db 654 FAENFOIKGSAVAVSRVSSIEGNWTVSNANAFGVVPMQNTICTRSDMTGLTTCOK 713
 QY 712 VDLITTKYINSIKPTQINSINLUDNATANYKGLATLNGNTVTLNHSQFTLSNNAFOIGN 771
 Db 714 VDLITTKYINSIKPTQINSINLUDNATANYKGLATLNGNTVTLNHSQFTLSNNAFOIGN 773
 QY 772 IRLSDNSTATVDNANNGVHLLTDSAPSLKNSHPHQIQKGTVTTLNNAFTWMBPST 831
 Db 774 IRLSDNSTATVDNANNGVHLLTDSAPSLKNSHPHQIQKGTVTTLNNAFTWMBPST 833
 QY 832 TQONTLNNSTITLNSAVSASNTPT--RRSLLETETPTSSAHRFTLTVNKLSGQGT 889
 Db 834 TQONTLNNSTITLNSAVSASNTPT--RRSLLETETPTSSAHRFTLTVNKLSGQGT 893
 QY 890 FQFTSLFQYSDKTLKSDNADGDTLISVNTGKEPTELOGLTVESKNDQPLSDTKAT 949
 Db 894 FQFTSLFQYSDKTLKSDNADGDTLISVNTGKEPTELOGLTVESKNDQPLSDTKAT 953
 QY 950 LENDVDAGALRYKLVKNDGEFLHNP-KGQELHNDLVABQABERTLEAKQVEPTAKOT 1009
 Db 954 LENDVDAGALRYKLVKNDGEFLHNP-KGQELHNDLVABQABERTLEAKQVEPTAKOT 1013
 QY 1010 GEPKVSRRARAAFPDPLPDQSLINALEAKQELTAFQSKAKTKYKVSRAV--FSD 1067
 Db 1014 GEPKVSRRARAAFPDPLPDQSLINALEAKQELTAFQSKAKTKYKVSRAV--FSD 1067
 QY 1068 PLIDSLFALBALEVTIDAPQSEKDLADQBEAK-QKQKDLISRSNALSLSATN 1126
 Db 1068 PLIDSLFALBALEVTIDAPQSEKDLADQBEAK-QKQKDLISRSNALSLSATN 1124
 QY 1127 SMLSYQDELDELFPDQASAVNTINADQKRRYSDFRAYSQOQKTNIRQISYOKALANR 1186
 Db 1125 SMLSYQDELDELFPDQASAVNTINADQKRRYSDFRAYSQOQKTNIRQISYOKALANR 1183
 QY 1187 ICAVFSHRSNDPTDQVKNHATLTMSGPAQIQWQDLQFGVNGTGISAKRAEVSQSK 1246
 Db 1184 ICAVFSHRSNDPTDQVKNHATLTMSGPAQIQWQDLQFGVNGTGISAKRAEVSQSK 1243
 QY 1247 HRKAINYGVASVQRLQGLQIOPYGVNRYFIEBNVQSEBYRVTPLANRYNGI 1306
 Db 1244 HRKAINYGVASVQRLQGLQIOPYGVNRYFIEBNVQSEBYRVTPLANRYNGI 1303
 QY 1307 RVDYTFPTDNIIVKYFFVNTVVDVSNANTVTLNVLVLOQFGRYQKXVGKALILHF 1366
 Db 1304 RVDYTFPTDNIIVKYFFVNTVVDVSNANTVTLNVLVLOQFGRYQKXVGKALILHF 1363
 QY 1367 QLSAFLSKSQSLGKQONVGVKLGXRM 1394
 Db 1364 QLSAFLSKSQSLGKQONVGVKLGXRM 1391

RESULT 9

US-10-080-505-11
 ; Sequence 11, Application US/10080505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RT/DCG/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080.505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 11

LENGTH: 1391
TYPE: PRT
ORGANISM: Faemophilus influenzae
US-10-630-505-11

Query Match 72.8%; Score 5297.5; DB 14; Length 1391;
Best Local Similarity 74.7%; Pred. No. 0;
Matches 1052; Conservative 109; Mismatches 216; Indels 31; Gaps 15;

QY 1 MKKTYFERNELTACISLIGVSOAMAGHTYGDYQYRPPANKGKTGAGAKITKYNQ 60
DB 1 MKKTYFERNELTACISLIGVSOAMAGHTYGDYQYRPPANKGKTGAGAKITKYNQ 60
QY 61 SOLVGTSMKAPMIDSVSNGVVALVBNQYVSAHVAVGTVDDGASGNPQCHRT 120
DB 61 SEMITGMKGVMPDSSMWKQGISLISEQHLSVAHVAGDVDDGMEGEPQCHRT 120
QY 121 YKIVKNNYKKNLHPYEDDYNPRLNFTYEAAPIDMTSNNGSTYSDRTKYPERVIG 180
DB 121 YKIVKNNYKKNLHPYEDDYNPRLNFTYEAAPIDMTSNNGSTYSDRTKYPERVIG 180
QY 121 YKIVKNNYKKNLHPYEDDYNPRLNFTYEAAPIDMTSNNGSTYSDRTKYPERVIG 180
DB 121 YKIVKNNYKKNLHPYEDDYNPRLNFTYEAAPIDMTSNNGSTYSDRTKYPERVIG 180
QY 181 SGQVWRNDQK-GQVAGAHYTLNNTNORAGNGSYLGGDYKAGYGLPLAG 238
DB 179 SGQVWRNDQK-GQVAGAHYTLNNTNORAGNGSYLGGDYKAGYGLPLAG 237
QY 239 SKDGSSEMTYDAEKOKMLINGILRENGPEKENGKPOLYKSYPDIEPRLTSLYT 298
DB 238 SKDGSSEMTYDAEKOKMLINGILRENGPEKENGKPOLYKSYPDIEPRLTSLYT 297
QY 299 RAGNGVYITSGNDG-----QGSITQSG-PSBIKITLANNSL-PLXKDKYHNPYDGP 352
DB 298 TNAEYFVIGSDHNGVATIKSTLPKALQPERIVGLYDNGSLDADKNGGESSYNGP 357
QY 353 NLYSPRLNNGELTYPMDOGSLIASEDINQAGGAYFENGFTYSPPSKQ-TQAGAHY 411
DB 358 NLYSPRLNNGELTYPMDOGSLIASEDINQAGGAYFENGFTYSPPSKQ-TQAGAHY 417
QY 412 SENSIVTWKVGVEHRLSKISGKTLHVOAGNGKKSISVBPQKVTLEQADQGNKQAF 471
DB 418 GRESTVEWQHNPBGRLSKISGLVNGKNGKGLSISVAKGLVLDQADESGOKAF 477
QY 472 SEIGLVSGRTVQINDKQFDTKYFPGFGLDINGSLTPKRIQNTDEGAMVNT 531
DB 478 KEVGIVSGRTVQINDKQFDTKYFPGFGLDINGSLTPKRIQNTDEGAMVNT 537
QY 532 TQAAVYITTESVLPNGNINQIDRKELIAYNGPFGTDKANKRLIYKPTEDR 591
DB 538 TQAAVYITTESVLPNGNINQIDRKELIAYNGPFGTDKANKRLIYKPTEDR 593
QY 592 TLLSGGTNLKGLITQKGLFFSGRPTPAHYHLNKRSEKGIPOGEIWNHDMKRT 651
DB 594 TLLSGGTNLKGLITQKGLFFSGRPTPAHYHLNKRSEKGIPOGEIWNHDMKRT 653
QY 652 FAENFOIKGSAVSRNYSIBGNVTVSNANATGVVPOQNTICTRSDMTGLTQOK 711
DB 654 FAENFOIKGSAVSRNYSIBGNVTVSNANATGVVPOQNTICTRSDMTGLTQOK 713
QY 712 VQLTITKINSIKPQIINGSINLTNATNAVKGIAKANTLTINHQFTLSNATQIG 771
DB 714 VQLTITKINSIKPQIINGSINLTNATNAVKGIAKANTLTINHQFTLSNATQIG 773
QY 772 IRLSDNSTAYVDNANNGVNLTDGACFLKNSHPSHOIQOQKGTVTLNANATMPBDT 831
DB 774 IRLSDNSTAYVDNANNGVNLTDGACFLKNSHPSHOIQOQKGTVTLNANATMPBDT 833
QY 832 TQNTLNNSTITLNSAYSSANNP-RRSLSTETTPTSAEHFNLTWNGKLSGGOT 889
DB 834 TQNTLNNSTITLNSAYSSANNP-RRSLSTETTPTSAEHFNLTWNGKLSGGOT 893
QY 890 FQFTSLEGYKSKLKLSDADGDTLLSVNTKKEPTELEQTLVESKKNQPSLKLFT 949
DB 894 FQFTSLEGYKSKLKLSDADGDTLLSVNTKKEPTELEQTLVESKKNQPSLKLFT 953

QY 950 LENDVDAGALEYKLVKNDGEPFLNPIKEDELHNDLVRAEOERTLEAKOPEPTAKTOT 1009
DB 954 LENDVDAGALEYKLVKNDGEPFLNPIKEDELHNDLVRAEOERTLEAKOPEPTAKTOT 1013
QY 1010 GEPKVRBRBAAPAPDPTLPDQSLNALBAKQELTAEQSKYAKKYKVRKRAV--ESD 1067
DB 1014 SKAVRBRBAAPAPDPTLPDQSLNALBAKQELTAEQSKYAKKYKVRKRAV--ESD 1067
QY 1068 FILDQSLFALFALAVIDAPQSEKORLAQEBAEK-ORQKDLISYSNALSLSATVN 1126
DB 1068 FILDQSLFALFALAVIDAPQSEKORLAQEBAEK-ORQKDLISYSNALSLSATVN 1124
QY 1127 SMLSVQDELRLFYDQASAVNTNINADRRYSDAPRAVQOKTLPOTIGORALNGR 1186
DB 1125 SMLSVQDELRLFYDQASAVNTNINADRRYSDAPRAVQOKTLPOTIGORALNGR 1183
QY 1187 IGAVFSSHSNDTPEQYKXNATLTMMSGFAQYQMGDLQGVVNGVIGSASMAEEGSRK 1246
DB 1184 IGAVFSSHSNDTPEQYKXNATLTMMSGFAQYQMGDLQGVVNGVIGSASMAEEGSRK 1243
QY 1247 IHRKALVGNASVQFRLQGLQIOPYGVNRYFERENYQSEEVRYKTSIAFNRYAGI 1306
DB 1244 IHRKALVGNASVQFRLQGLQIOPYGVNRYFERENYQSEEVRYKTSIAFNRYAGI 1303
QY 1307 RVDYFPTDNTSVKPYFVNVYDVSNANVOTVYVLOQPGRYWQKEVETAKELIHF 1366
DB 1304 RVDYFPTDNTSVKPYFVNVYDVSNANVOTVYVLOQPGRYWQKEVETAKELIHF 1363
QY 1367 QLSAFISKSQSGQLGKQOQNVKLGYYRW 1394
DB 1364 QLSAFISKSQSGQLGKQOQNVKLGYYRW 1391

RESULT 10
US-10-282-122A-65711
Sequence 65711, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haeelbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zykand Judih
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EUTRA-034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308

```

; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65711
; LENGTH: 1449
; TYPE: prt
; ORGANSIM: Neisseria meningitidis
; NS-10-262-112A-65711

```

Query Match	57.4%;	Score 4118;	DB 12;	Length 1449;
Best Local Similarity	56.2%;	Pred. No. 3.0e-285;		
Matches 816;	Conservative 214;	Mismatches 325;	Indels 98;	Gaps 17;

QY	10	ELACSGSIGVSAAMAHHYPPEDIDYOPRBAHKECTVAAONTVYVKKQCVGWSMT	63
Db	27	YLAICISFBIIPQAMAHHTYFGIANTYIOPRBAHVKGAFAKDIIVYVKKGBLVGWSMT	96
QY	70	KAPMIDFSVYSRNVAAALVENOYIVSVAAHVGYTVDVDFGAGZNPDRHFTYKIVKRNXY	129
Db	87	KAPMIDFSVYSRNVAAALVGDQYIVSVANHGXYNNVDFGAGZNPDRHFTYKIVKRNXY	146
QY	130	KXONLAFPEDDYXNPRILHFTYEAAPIDMTSNMNGSTYSDRTKYPERVATSSGQFWRND	189
Db	147	KPNSHPFPIVNGDIYHPRILHFTYDAPEVEMTSIDRNTYSDXKEKPERVATISGHHYKRYD	206
QY	190	QDKGQVAGAYHLLTGANTHQRQAGAGVSYLGGDVRAKAGEPTLEIASGKDGSGSPMT	249
Db	207	DDKHGDLTYSAGMLGGNTHMQGWMGNNGVYSLSGVVRANDYGPMEIIMAGDGSGSPMT	266
QY	250	YXAEKQKHLINGILBEGNPFBEKXKNGQLYTKSY-DEIFERDLHTSLYTAAGNYTTS	308
Db	267	YXCTNNKMLINGVLIQGYVPSGRNGEQILIKOMFYDDIYKRGTHYFVEERSNGHSPET	326
QY	309	GNDSGGSITQKSGISEIKITLANSL---PLAKEDYVNNPY--DGPNTYSRLNME	363
Db	327	SNNGGIVTETNEKTSNPKLVOTVLPFESLMETK--EPVYAAGVNDYRPLNME	384
QY	364	TYLPMQKQSSLLFASDINOAGGLYBEGNFTVSPNSQTMQAGLHVSNSYTWKXWG	423
Db	385	NLSFIDYGNKGLISNNINQAGGLYEPDFTVSPENNMQAGVHISDSYTWKXWG	444
QY	424	VHEDRLSTKIGGLTHVQAAGENKGSISVGGAKTILEQADQDNKQASLEGLVSGRTV	483
Db	445	VANDRLSTKIGGLTHVQAAGENKGSISVGGVTLILDQADKQKQAFSEGLVSGRTV	504
QY	484	QLNDDKQFDIKFEYFGRGRDLNGSLTFEKIONTDEGAMLVHNTTQAAVTTTGB	543
Db	505	QLNADNQNPNPKLYPFRGRDLNHSLSFRIQNTDEGAMLVHNAITTSYVTTTGB	564
QY	544	SYLPEGNNTKLYDKKELHYKMGFGTDRKNGRANLYKPTTEDRTLLSGGTNLKG	603
Db	565	STQPSGKINIRLNTSKELAYNGWFGKQKTYTNRANLYCPAABEDTLLISGGTNLKG	624
QY	604	DITQFKALFPSPRPPRAYVHNLKPKMSMEGIPQGEIYVMDHMINRFEKXNPQIKGS	663
Db	625	NIQVNGKLPFSRPPRIAHYNHLSGMSXQGEIPQGEIYVMDHMINRFEKXNPHIQSQ	684
QY	664	AVYSNVSSILEGNMTVSNNAANFTGVPPAQNTICTRSDTGLTTCQRYDLTPTYVNSI	723
Db	685	AVISSENVAKVEGMHLSHNAQAVFGVAFHQSHICTRSDTGLTNCVEKTTIDPKYIASL	744
QY	724	PKQIOWSINLTQNTAANVKGIAKLQNTVLTINHSQTLSSNATQIGH-----	771
Db	745	TKTIDISGVSLDHAHDLNLTGLATTNGNSANAGDRTYVSHNATQNGVLSYGNAAQATFN	804
QY	772	-----ILSSNSTAYVDANINAGVHLDTSAQFSLKNS	804
Db	805	QATLNGTSSASGNASPNLSNAAQNSGLTSPKAKAVHVSALNAGVSLADKXVHFENS	864
QY	805	HSHDIOQDGGCTVTLKAAWTYMSPTQLQNTLANSITLYASV-----SASNNY	866
Db	865	RFTQGLSGSSQDLNLEHKQSEVLPSPGELGNTLNDMTTILNSAVRHAAGAQOTGVSDT	924

QY	857	PP---RSLSLETTPTPSAHPNTLTYNKLSGGQGPQFTSLLEGYKSDKTLKLSNDAGD	913
Db	925	PPRRSRSSLSLSTPPIVSSESRFTLLTYNGLNQGIFFRPMELGYSRDKTLKLSASGCT	984
QY	914	YLLSVRNTGKEPPELTLEQLTIVSSKDNPLSDKLTKEFLNHDVDAGALRYKLVYNDGEFRL	973
Db	985	YTLANNTGMBPSSLIDQLTIVGKDNPLSLNNLFTLNQEHVDAAGARVQLIRKDGFEFL	1044
QY	974	HNPEKEBELNDIYRAE---GAER-----TLSEKQVEPTAKQT---GEYKYSRRARA	1022
Db	1045	HNPEVEQSLDKLGRKEAKKQAEKDNQASIDALIAAGDAAEKTESVAEP-----ARQ	1097
QY	1023	APPDTLDPSILNALAEKQAEILTAEQTSKAKTKIKYKRSKRAVFSDFPLQSLFALAELE	1082
Db	1098	AGEENV---GIMQAEHEKK-RVQADKDSALAKQREAFIRPACTAF-----RARRA	1144
QY	1083	VLDAPQOSEKDRLAGHEAKQ-RKQKLLISRSNSNLSESATYNSMLSYODELRLFLVD	1141
Db	1145	RRDLPQ-----PQVQPPQPPQFCRLISRYANSGLSSEFSATLNSYFAVQCHELDRFAE	1197
QY	1142	QAQSAVWTNLTADPKERYSDAFRAVYQOQKTNLRQIGVQKALAMGRIGAVSRSRSDNFD	1201
Db	1198	DRNNAVTSIGIRPETHYNSQDFRAVYRQ-TDLRQCMQXNGSRGVQLFSNNRLENFD	1256
QY	1202	EQYVNHATLTMSGPAQVQWGDLPQGNVTGTLSASNAEESQSKIHRAINYGNASYQ	1261
Db	1257	DGIGNSRLAHGAVPQVQIGIRFDIGISTAGPSSGSLSDQIGKLRBRRLHYGQAYRR	1316
QY	1262	FRLLQTLSDYPPGVNRYFERENYQSEEVYKPTPLAFNRNAGIRVDYTFPTNIEVX	1321
Db	1317	AGPGRGFEIEYIGAIRRYVQCADRYRNNMIALPGLAFNRIRACIDKIDSPFAOHISIT	1376
QY	1322	PFFVNTVDSNANVQTTNLTVLQDPGRVQKEVGKAEILHPQISAFISKQSGSLG	1381
Db	1377	PYLSLSTYDAAAGSKYRTVNTVAVLADQPGKTRSAFWNAEIKGFTLSLHAAAAGPQLE	1436
QY	1382	KQGNVGYKLGYSRR	1394
Db	1437	AKHSAGIKLGYSRR	1449

RESULT 11

US-1C-220-481-71

Sequence 71, Application US/10220481

Publication No. US20040110670A1

GENERAL INFORMATION:

APPLICANT: ARICO, Maria B., et al.

TITLE OF INVENTION: Heterologous Expression of Neisserial Proteins

FILE REFERENCE: CHIR-15883/010US

CURRENT APPLICATION NUMBER: US-/0/220,481

CURRENT FILING DATE: 2003-05-05

PRIOR APPLICATION NUMBER: PCT/IB01/00452

PRIOR FILING DATE: 2001-02-28

NUMBER OF SEQ ID NOS: 633

SOFTWARE: SeqMan9, version 1.02

SEQ ID NO 71

LENGTH: 1457

TYPE: PRT

ORGANISM: Neisseria meningitidis

US-10-220-481-71

Query Match

Best Local Similarity 56.2%; Score 4085; DB 16; Length 1457;

Matches 806; Conservative 219; Mismatches 328; Indels 110; Gaps 19;

QY	10	FLTKCSLIGVSAAGAGTFFGIDYQYRDPENKKGKFTGAGNICKYKNGOGSLVGSMT	69
Db	27	YIALCISFGLIQMAAGHTTFGLINTQYTRDPRNKGKFAVGAQDIEYNNKGELVGSMT	86
QY	70	KAPMIDPVSNSNGAALVENQYIVSVAHNVGYTDVDFAGANNPQDHRFTYKIVKNNY	129
Db	87	KAPMIDPVSNSNGAALVDQYIVSVAHNGGYNNVDPGAGENNPPQDHRFTYKIVKNNY	146

Gy 70 KAPMIDFSVSRNGVAALVENQYISVAHNQYETVDPGAGGNNDQHRFYIKIYKRNRY 129
| | | | | : | | | | | : | | | | |
Db 87 KAPMIDFSVSRNGVAALVGQYISVAHNGYNVDFGAEGRNDDQHRFYIKIYKRNRY 146

QY	232	LHSLYTRGKGVWYVSGNNQGGSLTQK---	SGIPSEIK---	ITLANSI,LEKXKOKYH	345
Db	229	THSHFTRPQNKXISFMDNKGCTGKNAKHEHNSFNPKRKTQVLEPNVSLSETRAPVY	358		
QY	346	NPRYDGNITYSPRLNNGETLYFMQDKOSLIFASDINQAGALYFEGNLFVSPNSQCTWO	405		
Db	359	FAA--GGVNSYRPLANNENISFIDGKGEJLITLSININGAGALIFQSDPTVSPENNETWQ	417		
QY	406	ANGIHNSXNSTYTKVNVGVSHDRBSIKCKGLTAYQAKGNKGSISVQDGKVTLEQADQ	465		
Db	418	GAGVHISSESTVTKVNVGAVANDRLSKICKGLTAYQAKGNQGSISVQDGTVILDDQADK	477		
QY	466	GKQAFSEIGLVSGRGVQANDKOPDDIKYGGPAGRLDNLNHSITFKRLOMTDGGAM	525		
Db	478	GKQAFSEIGLVSGRGVQANDKOPNDKLYFPGFGRGLDNHSISLSPKRLQNTDGGAM	537		
QY	526	IVNHNTQANVTITNGESIVLPNGNNINKLDYKEIAYNGMFGETDNKNGHRLNLTLYK	585		
Db	538	IVNHNDKESTVITNGKDA--ITGNN--NSIDSKKEIAYNGMFGKOTTKMGRLNVYQ	595		
QY	586	PTTEDTLLSGTNMGSD--TCTKGKLPFSGRTPPHAVNHLNKRMSMGISQOSIYWDH	645		
Db	596	PAADRLTLLSGTNMGK--TQTNKSLPFSSRTPPHAVNHLNDMSKXGIPROGIYWDV	655		
QY	646	DMIRTFEKAENQJGGSAVWSRNVSIIZCANVTVSNANATFEGVAVPOQNTLCRSPWTS	705		
Db	656	DMIRTFEKAENQJGGQAVVSRNAVKKXDDMLSHQAVFGVAPROSHITLCRSPWTS	715		
QY	706	LTCQAVDLDTPKYINSIPTQOINGSINLTDNATANVGLAKNGVNTLTHNSQPTLSKN	765		
Db	716	LTCQAVCKLTDPKVIASTIKTDISQVLDADHAFMLTGLATLTMKSLANGDTPTVYSHH	775		
QY	766	ATQIOLK-----IRSDNSTKYTCYMAA	786		
Db	776	ATQKSLISIVKGAATNTQATLNGTSSAGMSFNLSPHAVQNSLITSGNKAKVASHS	835		
QY	787	LNENVALTDSQFSLKNSHFHQJQGGKTVTLLENATVTPSPDTLLQNTLITANSSTILAN	846		
Db	836	LNQVSLADKAVFHFESRTEQJISGGKDTALHKOEMTLPSSTELGNLMDNATITLN	895		
QY	847	SAY-----SASNNTR-----RSLSETTPTAHERPVLTVNGKISQCGTPOF	892		
Db	896	SARHDAQAQGTGATAPRRSRRSRSLSTVPTVESSEFNLTVNGKLNQCGTFR	955		
QY	893	TSSLPFYKSDKLTISNDAEGYILSVRTEKEPTELEQVLYESKQNOPLSDPKFPLEN	952		
Db	956	MSELFEYRBDKILABSEBQYTLAVNTGNESASISQOVTYBESKONKRLSPNLTPLQN	1015		
QY	953	DHYDAGALRYKLVNNDGEPLHNDPKQOELHNDVPAE--QAE-----TLEAKQVE	1002		
Db	1016	EHWDAQAWMYOLIKQSEFFLNPVYKQESLDTGLAKAKKQAEKQNASIDALILAGSD	1075		
QY	1003	PIKTKQT--GEPRVSRBARAARAPPTLPDQSLNLTLEKQELTAETQKSYAKTKYKXK	1061		
Db	1076	AVKETSVAEP-----ARQGENV--GIMQAEERK--RQADKOTLAKAKREAFER	1122		
QY	1062	RAVSDPILLOSLFALAEALVITAPQOSSEKRLQAEBAEKOROKOLISRYNSALSSEI	1121		
Db	1125	PATTAEP-----RARRARQDLPOLQPOPOPOPORDILISRYNSGLSEF	1167		
QY	1122	SATVNSMLSYODELDLPVDOQASVNTMLAQQKRYNSDAPFAAYQOQCTMLRDQIQYQA	1181		
Db	1168	SATLNSVPAVODELDTRFADBRNNVMTSGIRPTGHVNSQDPAFYAQQ--TDLRQIQYQAN	1226		
QY	1182	LANGRLGAVFSSHRSNDTDFEQVKNHNTLTMNSGAQYQOWMDQJQFVNVNGTGISASMAE	1247		
Db	1227	LGSGRGVGLFSHNTEFTDDGINSARLAKHAGVFGQGDIDREYJISAGAGFSSGSDS	1286		
QY	1242	EQGRKIKHKAIVNVAASQVRLQGLQIQYQVNRKRYTERENVYQSEERYKVTQSIAFR	1301		
Db	1287	GIGCKIKRRVVLHSGQARRPAGGFGFIEPHGATRRYVQADAYRRNNNTATATGATFR	1346		

```

01 1302 YAGIVGVETTFPPIDINSVKPEFPFNVDVSNANQTVNLVTIQDPERFVWQGEVETLA 1303
02 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
03 1347 YRAGIADISREDAQHSITPPLSLSTIDMAASKRFRVNVANVAADPSKTRSAKSVNA 1402
04 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
05 1362 EILAFQISAFISKSQSGSOLGAKQGVNWKLGCVN 1394
06 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
07 1407 EIKGFPLSLHMAAKGPGQLAQHSAGIKLGCVN 1439
08
09 RESULT 14
10 US-09-839-996-3
11 ; Sequence 3, Application US/09839996
12 ; Publication No. US20030009010A1
13 ;
14 GENERAL INFORMATION:
15 APPLICANT: St. Gene III, Joseph W.
16 Falkow, Stanley
17 TITLE OF INVENTION: Haemophilus Adherence and Penetration
18 Protein
19
20 NUMBER OF SEQUENCES: 9
21
22 CORRESPONDENCE ADDRESSES:
23 ADDRESSER: Flehr, Hombach, Test, Albritton & Herbert
24 STREET: 4 Embarradero Center, Suite 3400
25 CITY: San Francisco
26 STATE: California
27 COUNTRY: United States
28 ZIP: 94111-4187
29
30 COMPUTER READABLE FORM:
31 MEDIUM TYPE: Floppy disk
32 COMPUTER: IBM PC compatible
33 OPERATING SYSTEM: PC-DOS/MS-DOS
34 SOFTWARE: PatentIn Release #1.0, Version #1.25
35
36 CURRENT APPLICATION DATA:
37 APPLICATION NUMBER: US/09/839,996
38 FILING DATE: 20-Apr-2001
39 CLASSIFICATION: <unknown>
40
41 PRIOR APPLICATION DATA:
42 APPLICATION NUMBER: US/08/296,791
43 FILING DATE: 25-AUG-1994
44
45 ATTORNEY/AGENT INFORMATION:
46 NAME: Treacartin, Richard F.
47 REGISTRATION NUMBER: 31,801
48 REFERENCE/DOCKET NUMBER: A-59941/RET/RMS
49
50 TELECOMMUNICATION INFORMATION:
51 TELEPHONE: (415) 781-1989
52 TELEFAX: (415) 398-3245
53 TELEX: 910 277299
54
55 INFORMATION FOR SEQ ID NO: 3:
56 SEQUENCE CHARACTERISTICS:
57 LENGTH: 1541 amino acids
58 TYPE: amino acid
59 TOPOLOGY: unknown
60
61 US-09-839-996-3
62
63 Query Match 23.0%; Score 1673.5; DB 10; Length 1541;
64 Best Local Similarity 28.5%; Pred. No. 1.8e-108;
65 Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;
66
67 QY 6 FLNLPLTACISLGSQAAMGHTYFGIDYQYDRFAENKKEFTVGAQINIKYKNGQGLVG 65
68 | F I N I A L T A V A L T E Y T A A L V R D D V D Y Q I P R D F A E N K K F S V G A T V A L V K D K N K D I G 65
69 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
70 Db 6 TALPNSGIMDESVVDVDRKRLATLINFQYVAGVKSIVNSGVSEIHFQNLNNGMNNGNAKAH 125
71 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
72 QY 66 TSWTETK-EMIDESVVS-RUGVAALVENVQYIVSVAH-NVGYTDVDRG-AEGV---NPDDH 117
73 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
74 Db 66 TALPNSGIMDESVVDVDRKRLATLINFQYVAGVKSIVNSGVSEIHFQNLNNGMNNGNAKAH 125
75 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
76 QY 113 RFT-----YKIVKNNY-----KKDMLHPYEDDYHNRFLHKEFTVEAAPID-MTS 160
77 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
78 Db 126 RDSVSEKRYSVSEKNEYTKNLNGKRTVTTEDDTQYKREDDLYMPRLDKVTEVAPIASFA 185
79 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
80 QY 161 NMAGSYSDR-KTKPKRYVIRSGSGRQPFNRNDQKGD-----YAGAYHTL 203
81 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
82 186 ISSAGTYNDQKRPDAVIRVSGSGQF---IYKGNYSLLINHEVGNGNNLKTLDGDAYTGG 242

```

QY 204 TAGTHNGAGAGVSYL-----GVPRAGEYGP.LPAGSGDGSMPYTAERKQK 257
 Db 243 IAGTYKVNENKGLIGNSKEHSDPFGILISQDPLNVAVLGDGSDLPFYRERKMK 302
 QY 258 LINGILREAGPEEGE-NGFOLVAKSYFDEIFERDILTSL-----YTAGNG-VYTIIS 309
 Db 303 LFSGSYDFWAGYKKSQWQEMWYYSQPTQDVLMKMSAGSLGKTDYMSNSGKCTITTY 362
 QY 310 NDNGQSSITQKSGITSEIKITTLANGLPLKEKQKHYRPRYDGENIYSPALNGETLYEMD 369
 Db 363 GE-----KSLNDVLAD-----GDKP-----XHGKSYTF-- 386
 QY 370 QKQSLIFASDINGAGGLYPEGNFTYSPNS-NTWQAGIHVSENGTYWKNGVGEHPR 428
 Db 387 EGSETLTNNIDQAGGLFEBGYEKGTSQNTYWKAGAVSAEKGTYWKVHPQYDR 446
 QY 429 LSKIGKTLHQAQENKSGISVQDKVLEQOADDQONKASEIGLWGRGVQOLND 488
 Db 447 LAKIGKTLIVEGTGDKNGSLKVGDG-VILKQINSGG-QHAFASGIVSGRSTVLVND 505
 QY 449 KQPDTEKTFYFGRGRDLNNGSHLITKRLQNTBQAMTVNHTTQANVTITGESITVP 548
 Db 506 KQVDPNSIYFGRGRDLNNGSHLITKRLQNTBQAMTVNHTTQANVTITGESITVP 565
 QY 549 N----- 549
 Db 566 NLTPEYNIAPDEPNPAPFRIRKQGGQLYALNENTYTAARKASTREBELFRANSGBENEN 625
 QY 550 -----GNINKLYRKEIANGVMEGETDKKNGRNLNLYRPTEDRTLLSG 597
 Db 626 MLYKGTSEAKRNMNMNNHNNERMNGFNGYFEB-BEKKNGNLMNTPKSGEQRLILNG 684
 QY 598 GTNLKGLITQTKGLFBSGRPTPHANHLAKRSGSMGT-----PQGETVWDHNM 647
 Db 665 GTNLKGLITVEKGLTFLSGRPTPHA-----SDIAGISTQWQPHRAENNEVEWEDWM 736
 QY 648 INETFAENFO:KGSAYVS-RVWSSIEGNTVYSNANAPFGVWPNQNTCTRSMTG 706
 Db 737 INNFKATIMNTGASISGRVNAITNTTASNAQVHIGY-KTGDTVCVASDITG 794
 QY 707 TTQKQVLDITVINSIPKTIQINGSINTLNATNANVGLAKNENVTLLNHSQPTLSNNA 766
 Db 795 VTCITDKLSD-KALNSFNPTNLRGVNLTESNMF-VLGRKANLEGTIGRNGSQ----- 845
 QY 767 TQIGNISDINSTA-VUNANLNGNMTLDSQFSLKNSHESQIQDGKTTVILENMT 826
 Db 846 -----VLTENS-----HMLTGN-----SOVQOLDLANGH----- 871
 QY 827 MFSDTTLQNLILNKSITLINSAYSASNNTPRRSLTETTPSAHEHFNLLYNGKLSG 886
 Db 872 -----IHNSA-----DNSNNVT-----KNTLITVN-SLSG 896
 QY 887 QGTROPSTSLFSGYSKDKLSDNDAGDYLLSVNTGKEPETLEQULTVESKQNPISDXL 946
 Db 897 NGSPFYLLDLSNKQDQKVVYTKSATGNFTLVADKTEBPNH-NELTLEDAKQ--RDHL 953
 QY 947 KETLENDHVDAGALRYKLYKNQGEPRINPEIKEOHLNDLY----- 987
 Db 954 NVSLVGNVVDLAGMKTCLRNVRNGERYDLYN-EVEKKNQVVTNITTPNNIQADVSPV 1011
 QY 988 -----RAQOERTLEAKOVEP----- 1004
 Db 1012 SNNEIARVDAVPVPPAPATPSETTEYVANSKQBSKIVKKNQDAETTMQNRVAK 1071
 QY 1005 -----AKTQGE-----PKVSRRAA 1020
 Db 1072 AKSNVANTQUNEVAOSGETKETQTEKETAIVEKEKAYETEKTEQVFKTQSYVP 1131
 QY 1021 RAAPETLPGOS-----LNLALAK-CALTAETQSKAKTKVNSKRAVSPBLD 1071
 Db 1132 KQBSSTVQPOAFAPARENDPTVNIKPSQNTTADTBQAKETS-----SNVEQPVTE 1185

QY 1072 QSLFALPALEVTIDAPQ-----SEKRIAOEAEKQKQK----- 1107
 Db 1186 ST-----TVNTGNSVVENENTTPATOPTVNSSESNKKNHRSVSVPHNVEPATSSND 1243
 QY 1108 -----CLISRYNSAULSELSA-----TNSMLSYQDELRILV-DOASAWMTIAD 1154
 Db 1244 KSTVALCOLTSTVNTNVALSDAPAKAQVALNVRKAVASOHLISQLENNNEGYNNWVSNTSM 1303
 QY 1155 KERYSDAFRAYOQOQNTNRLQIGVOKALANG-RIGAVFHSRSDNTDEQVKNHATLTVM 1213
 Db 1334 MNVSSSYRRFSPSKTQ-TQLMDQITLNNVQDGGVTFYRNSNNNDKATSKN-TLQAV 1361
 QY 1214 SGPAQY-----QKQDLOFVNVGTGILASRAEBSQKILHRAKAINGVNAYSOFRLQGLT 1269
 Db 1362 NPYKTYVADNMW-----YIGIDLGYGFKOSKLTQNNNAKFAHHAQFGHLPAGKAPNLNFGI 1418
 QY 1270 QPYFVNVFIERENYQSEEVAVKTPSLANRYNAGIRVDTFPTON-SVKRYFPVNVY 1329
 Db 1419 TPIVGRVSTLSNADALDQARIKVNPISVKTAFAQVDLSTY-HLGEFVTLISARY- 1476
 QY 1330 DVSANNOQTVNLTVALQOPFRYQKEVGLKAEILHFOISAFISKSQSQGLQKQWVGV 1389
 Db 1477 DANQSGKINWGEYDPAVNVENQQYVAGKJLYHNVKLSLGGULTAKQAEKQTAELX 1536
 QY 1390 LGY 1392
 Db 1537 LSF 1539

RESULT 15
 US-10-645-655-3
 ; Sequence 3, Application US/10645655
 ; Publication No. US20040063908A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene III, Joseph W.
 ; Falkow, Stanley
 ; TITLE OF INVENTION: Haemophilus Adherence and Penetration
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/645,655
 ; FILING DATE: 20-Aug-2003
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/98/296,791
 ; FILING DATE: 25-Aug-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Treccartin, Richard F.
 ; REGISTRATION NUMBER: 31,801
 ; REFERENCE/DOCKET NUMBER: A-59941/RET/RMS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1541 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-10-645-655-3

Query March 23 04; Score 1673.5; DB 12; Length 1541;
 Best Local Similarity 28.58; Pred. No. 1.6e-108;
 Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

QY 6 FELNELTACISLIGVSOAMAGHTYDIDYDDEPAENKCKTVOAONIVYKQOOLVG 65
 DB 6 FKLNFALTAVALTETPAALVDDVYQIFRDEPAENKCKSVGATNIVAKOKKNDIG 65
 QY 66 TSMYKA-EMIDESVVS-RNGVAALVENOYIVSAH-NVGYTDVDFG-ABGN---NPDOH 117
 DB 66 TALPNGIEMIDFVSVDVDRKIATILNIPQYVGVAVSNVSELAHGNLKNMNNNGAAKH 125
 QY 118 EFT-----YKIVKANN-----KCONHPYEDYANPRHAKFTTEAALPD-MTS 160
 DB 126 FOVSSEENRYSVEKKEYPYTKLNGKVTYTBEOQOKREDDYVPRLDKFEVAPLEASTA 185
 QY 161 NNNGSTYSDRTKYPERVIRIGSGRQFMRDODKDO-----VAGAYAYL 203
 DB 186 SSDACTYNDOKPIPAFRLSGSSQF---IYKGNNTSLINNHEVGNLKLVDHATYTG 242
 QY 204 TANTHNOAGANGSYTLG-----GEVRAKAGEYGPLIAGSKDSDGSPMEIYDAEKQW 257
 DB 243 IAGTYKVNHENNGI.GFGNSKEHSDPKGLISQDPLTYAVVLGDSGSEIFYDAEKQW 302
 QY 258 LINGLIEBGNFEGKE-NQFQVLRKSYDELFERDJAFTSL-----YTRANG-YVTITSG 309
 DB 303 LFLSSYDFWAGVKKKMOEMNLYKQFVKVILKDKDSAGSLIGSKTIDSWMSNCKSTITIG 362
 QY 310 NNNGGSIYOKSGIPSEIKITLANNSLPLKENDKXNPRYDGNIVYSPFLNNGETILYMD 369
 DB 363 GE-----KSLNVDLAD-----GKQKP-----NHKSVTP-- 386
 QY 370 QKQSLIFASDINQAGAGLYFEGNFTVSPNS-NQTMQAGLHNSENVTWKXNGVEHR 428
 DB 387 EESGLTNANNIDQAGAGLFFEGDYEVKTSNWTWKAGVSVAEKGLVTKVHPQDRA 446
 QY 429 LSKIGKGLIYAKAGNKSISVGGKYLLEQADQOKKQAFSEIYLSGKGTQLAND 488
 DB 447 LAKIGKGLIYEGTGNKSLKVGAGTILKQDNGSG-QAHAFSAVSIAGSGSTIYVLD 505
 QY 489 KQPDYKFEYEGRGHLDLNGHS-TEFKRIQNTDEGAMTYNNTQANVITIGNSIYLP 548
 DB 506 KQVDNNSIYEGPRGGLD.LNGNSLTFPHITNDQAKLYNRMNTNNSITITIBSLITDP 565
 QY 549 N----- 549
 DB 566 NITPYNIDAPEDENPYAFERIKDGGQLYLNTENYVYALKKCASTRSELPRKNGSENEN 625
 QY 550 -----GNINKLDYKEKELAYNGFEGTDKNKNGRLMLYKPTEDRTLLISG 597
 DB 626 WLNGKTSDEAKKNVNMHLNNEKNGNGYFGE-EEGKKNQNLWVTEKGSQONRFLITG 684
 QY 598 GTNLKGLITQTKGLFFSGRPTTPAHVNLNKRSEMEGI-----POGEIYWDHW 647
 DB 685 GTNLNGDLIYKGLTFLSGRPTPA-----RDIAGISSTKDPHFAENNEVVEDW 736
 QY 648 IRRFTKAKNTQIKGSAVVS-RNVSSIEGNVTVSNANATFVVPNOQNTICTRSMDTGL 706
 DB 737 IRRNKATVMTVGNMSTLSYGRVANITSNITASNAAQVHIGY--KTDGVTCVRSDYTG 794
 QY 707 TTCCQVULDTQYKINSIPKTOJNGSINLTQNAIYAVKGLATNGVNTLNHSQPTLSNNA 766
 DB 795 VTCTTDKLSD-KALNSENFTNLRGNVLTESANF-VLQKANTLFGTIOGRNSQ----- 845
 QY 767 TQIGIRLNSNTATVDMAN.LNGVNL.TDSAQFSLKNSHFSHQDQKGTVTLNENATWT 826
 DB 846 -----VALTENS-----HMHLCGN---SDVHQDLANGH----- 871
 QY 827 MESPDTLQNTLNNSTIYNSAYSSANNTPRRRSLETETPTTSABHRNLTLYNGKLSG 886
 DB 872 -----IHLNSA--DNSNNVT-----KXVTLTVN-SLSG 896

QY 887 QGTQFTSLIFGKSDKLSNDABGDYILSTRANGKEPPELTQTLVSKNOPLSKL 946
 DB 897 NGSEFYIYIDLNSKQDDKXVVKSAQNFTLOVADCTGEPNH-NEILTLPASRQO--RDHL 953
 QY 947 KFTLENDVDAKALRYKLVKNDEPFLRHPPIKEOELHNDV----- 987
 DB 994 NVSLVGNVYDLAKMKYKLRNVNGRIDLYNP--EYKRRQVTDNTITTPNNIQAQDVSPV 1011
 QY 988 -----RABQAEKRLAEKQVEPT----- 1004
 DB 1012 SNNEIARVDEAPVPAPATPSETTETVAENSKQESITVSKNEQATETTAQNEVAK 1071
 QY 1005 -----AKTQGE-----EYKSRRA 1023
 DB 1072 AKSNVKAQNTQIMEVAGSSEKTEQTTEKATATEKEKAKVETKEQVEKVSQVSP 1131
 QY 1021 RAAPOTLPDS-----LNNALAEK-QAEITAEQKSAKTKYKRRKRAVFDPLD 1071
 DB 1132 KOQSEIYQOPAPRENDPTVNIKEPQSQNTNTVADYBOPAKETS-----SNEQPVTE 1185
 QY 1072 QSLFALAEALVETAPQO-----SEKDLAQSEAEKQKQK----- 1107
 DB 1186 ST-TVNTGNSVVENPEVTPPATPTOPTVNSSSNKKPKRRRBSRVSPHNVPATSSND 1243
 QY 1108 -----DLISRYSNLSALSELISA-----TVNSMLSVQDELDRLPV-DQAQSAVWTNIAOD 1154
 DB 1244 RSTVALCDLITNTNNAVILSDARAKAQFVALNVKAVSQHLSQLEBMNNEQVYVWVWSITSM 1303
 QY 1155 KRRYSDAPFRAYQOQKTNLRQIGVOKALANG-RIGAVPSSRSQNTPEQVKNHATLTM 1213
 DB 1304 NKVYSSQTRRBSKSTV-TQLGMDQTTISNNVQGLQFTYKRSNNPDAITSKA-TLAQV 1361
 QY 1214 SGFAQY-----QWGD-QQFQVNVGTGISASPAECSRKIHRKALNYVNASYQFRLQOLGI 1269
 DB 1362 NFYSKYVADNMH---YGIDIGYKFGQSLQTNHNAKPAHNTAQFGLTGAFAFLNGRGI 1418
 QY 1270 QPITGNRFFIERKRYQSEEVRYKTSLSAENRYAAGRVDITTPPDNISVYEPFYNAVY 1329
 DB 1419 TPIVGRVYSYLSNADPFLADQARIKVNPISVKTAFAQVDSLYT-HUGFSVYPIISARY- 1476
 QY 1330 PVSNAVQTTVNLTVLQOPFGRYQKEVGLKAEILHFOISAFISKQSGSLQKQOONGVK 1389
 DB 1477 DAQSGSKLVNGDYFRATYENQOQVNGKLKTHNVKLSLISGLTKAKQAEKQKTAELK 1536
 QY 1390 LGY 1392
 DB 1537 LSF 1539

Search completed: June 23, 2004, 17:39:00
 Job time : 67 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2004, 17:30:13 ; Search time 195 Seconds

(without alignments)
6977.532 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273
Sequence: 1 MKGVPRINFLINACISLGIV.....SQSQQLKQCNVGVKLGRM 1394

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 6019581 seqs, 976035377 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Pending Parents: AA Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
28: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
29: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
30: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
31: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
32: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
33: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
------------	-------------	--------------	----	-------------

1	7273	100.0	1394	1	PCT-US03-05226-2	Sequence 2, Appli
2	7273	100.0	1394	31	US-10-645-655-2	Sequence 2, Appli
3	7273	100.0	1394	31	US-10-687-046-2	Sequence 2, Appli
4	7246.5	99.6	1395	1	PCT-US03-05226-7	Sequence 7, Appli
5	7246.5	99.6	1395	31	US-10-687-046-7	Sequence 7, Appli
6	6115	84.1	1434	1	PCT-US03-05226-9	Sequence 9, Appli
7	6115	84.1	1434	31	US-10-687-046-9	Sequence 9, Appli
8	5827.5	80.1	1411	1	PCT-US03-05226-17	Sequence 17, Appli
9	5827.5	80.1	1411	31	US-10-687-046-17	Sequence 17, Appli
10	5807	79.8	1436	1	PCT-US03-05226-13	Sequence 13, Appli
11	5807	79.8	1436	31	US-10-687-046-13	Sequence 13, Appli
12	5303.5	72.9	1391	1	PCT-US03-05226-15	Sequence 15, Appli
13	5303.5	72.9	1391	31	US-10-687-046-15	Sequence 15, Appli
14	5297.5	72.8	1391	1	PCT-US03-05226-11	Sequence 11, Appli
15	5297.5	72.8	1391	31	US-10-687-046-11	Sequence 11, Appli
16	4178	57.4	1449	22	US-09-791-537-115250	Sequence 115250, A
17	4178	57.4	1449	28	US-10-282-122A-65711	Sequence 65711, A
18	4090.5	56.2	1454	20	US-09-673-896-2	Sequence 2, Appli
19	4090.5	56.2	1454	32	US-10-742-345-2	Sequence 2, Appli
20	4085	56.2	1457	17	US-09-303-518D-650	Sequence 650, App
21	4085	56.2	1457	22	US-09-791-537-40489	Sequence 40489, A
22	4085	56.2	1457	23	US-09-806-866-87	Sequence 87, Ap
23	4085	56.2	1457	23	US-09-806-866A-1047	Sequence 1047, Ap
24	4085	56.2	1457	26	US-10-018-470A-87	Sequence 87, Appli
25	4085	56.2	1457	28	US-10-742-345-4	Sequence 4, Appli
26	4083	56.1	1457	20	US-09-673-896-4	Sequence 4, Appli
27	4083	56.1	1457	32	US-10-742-345-4	Sequence 4, Appli
28	4072.5	56.0	1457	22	US-09-791-537-110808	Sequence 65027, A
29	4072.5	56.0	1457	22	US-09-791-537-110808	Sequence 110808, A
30	4067	55.9	1449	17	US-09-303-518D-652	Sequence 652, App
31	4061.5	55.8	1449	17	US-09-303-518D-654	Sequence 654, App
32	4046	55.6	1455	30	US-10-484-703-32	Sequence 32, Appli
33	4044.5	55.6	1459	28	US-10-220-481-73	Sequence 73, Appli
34	3409.5	46.9	1178	30	US-10-484-703-34	Sequence 34, Appli
35	3352.5	46.1	1136	30	US-10-484-703-36	Sequence 36, Appli
36	2941	40.4	956	30	US-10-484-703-33	Sequence 33, Appli
37	2886	39.7	914	30	US-10-484-703-35	Sequence 35, Appli
38	2347	32.3	979	17	US-09-303-518D-648	Sequence 648, App
39	1676	23.0	1431	20	US-09-674-546A-2601	Sequence 2601, Ap
40	1673.5	23.0	1540	9	US-08-578-788-2	Sequence 2, Appli
41	1673.5	23.0	1541	1	PCT-US03-05226-3	Sequence 3, Appli
42	1673.5	23.0	1541	31	US-10-645-655-3	Sequence 3, Appli
43	1673.5	23.0	1541	31	US-10-687-046-3	Sequence 3, Appli
44	1662.5	22.9	1545	1	PCT-US03-05226-4	Sequence 4, Appli
45	1662.5	22.9	1545	31	US-10-645-655-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
PCT-US03-05226-2
Sequence 2, Appli
GENERAL INFORMATION:
APPLICANT: St. Gene, Joseph W
TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
FILE REFERENCE: A-59941-1/RT/DCR/DHR
CURRENT APPLICATION NUMBER: PCT/US03/05226
CURRENT FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: US 08/296,791
PRIOR FILING DATE: 1994-10-25
PRIOR APPLICATION NUMBER: US 09/839,996
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 1394
TYPE: PRT
ORGANISM: Haemophilus influenzae
PCT-US03-05226-2
Query Match 100.0%; Score 7273; DB 1; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1394;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKTYFRNLPLTACISLIGIVSOAMAGHTYGDYDYYRPAENKGTGVAQONIKYKQ	60	
DB	1	MKTYFRNLPLTACISLIGIVSOAMAGHTYGDYDYYRPAENKGTGVAQONIKYKQ	60	
QY	61	GQVGTSTKTAPMIDPSVSVSNGVAALVENQYIVSVANVGYTDVDFGABNNPDQHEFT	120	
DB	61	GQVGTSTKTAPMIDPSVSVSNGVAALVENQYIVSVANVGYTDVDFGABNNPDQHEFT	120	
QY	121	YKTYKRNKYKKNLHYEDDYNHRLKRYTTEAFLDWTSMNSNSTSTRTYPRRPAIG	180	
DB	121	YKTYKRNKYKKNLHYEDDYNHRLKRYTTEAFLDWTSMNSNSTSTRTYPRRPAIG	180	
QY	181	SGROPFRNDQKQJYAGAXHYLTAGNTHQKAGNGSYLGQVRAAGYGLPIASK	240	
DB	181	SGROPFRNDQKQJYAGAXHYLTAGNTHQKAGNGSYLGQVRAAGYGLPIASK	240	
QY	241	GDSSGPMFYDAEKQKWL-MGILBGNPFEGKNGFQVKKSYEDBIFERDLHTSLYTRA	300	
DB	241	GDSSGPMFYDAEKQKWL-MGILBGNPFEGKNGFQVKKSYEDBIFERDLHTSLYTRA	300	
QY	301	GNGVYTTISGNDNGQSGITQKSGISEIKITLAMSPLKEXKQKHNFRIDGPRVISPAN	360	
DB	301	GNGVYTTISGNDNGQSGITQKSGISEIKITLAMSPLKEXKQKHNFRIDGPRVISPAN	360	
QY	361	NGEITYEYDQKQSLIFASDINOAGAGLYEGNFTVSPNSQVQAGIHVSENSTYWK	420	
DB	361	NGEITYEYDQKQSLIFASDINOAGAGLYEGNFTVSPNSQVQAGIHVSENSTYWK	420	
QY	421	VNGVEHRLSKIGKTLHYQAKENGKGSISVSGKTYLEQADDQKQKAFSEISLVSGR	480	
DB	421	VNGVEHRLSKIGKTLHYQAKENGKGSISVSGKTYLEQADDQKQKAFSEISLVSGR	480	
QY	481	GTVOANDKQDPTDKYFEGRGGRLJLNGSLIFKELONTGEMITVHNTTOANVIT	540	
DB	481	GTVOANDKQDPTDKYFEGRGGRLJLNGSLIFKELONTGEMITVHNTTOANVIT	540	
QY	541	GNSIVLPNGNINIKLDYKELAVNGVGETDKNGKNGRLNLYKPTTEORTLLSGGTN	600	
DB	541	GNSIVLPNGNINIKLDYKELAVNGVGETDKNGKNGRLNLYKPTTEORTLLSGGTN	600	
QY	601	LKGDITQTKKGLPFSSRPPEHYNHAKKSEMEGIPQETVYMHQIMRTKAKNPOIK	660	
DB	601	LKGDITQTKKGLPFSSRPPEHYNHAKKSEMEGIPQETVYMHQIMRTKAKNPOIK	660	
QY	661	GGSAYSVANSISIEKMTVSNNAATGQVVPNOQNTICTRSPWTGLTTCQKVDLDTYVI	720	
DB	661	GGSAYSVANSISIEKMTVSNNAATGQVVPNOQNTICTRSPWTGLTTCQKVDLDTYVI	720	
QY	721	NSIPKQINSINLTNATATANKGLAKNGVTLTNHSGFTLSNNAQOIGNRLSDNSTA	780	
DB	721	NSIPKQINSINLTNATATANKGLAKNGVTLTNHSGFTLSNNAQOIGNRLSDNSTA	780	
QY	781	TYDANALNGNHLTDSQSLKNSHFSHQIQSKGTYYTLENAITWBSSTLYQULTAN	840	
DB	781	TYDANALNGNHLTDSQSLKNSHFSHQIQSKGTYYTLENAITWBSSTLYQULTAN	840	
QY	841	STTLTNSAYSSASNNTPRRRSLETCTPTSAERFNTLVNGKLSGGGTQFQFSSIFGX	900	
DB	841	STTLTNSAYSSASNNTPRRRSLETCTPTSAERFNTLVNGKLSGGGTQFQFSSIFGX	900	
QY	901	SDKLSNDABGYILSVN-TSEPETLEQTLVESKDNQPLSDKLFYTLNDHYDAGL	960	
DB	901	SDKLSNDABGYILSVN-TSEPETLEQTLVESKDNQPLSDKLFYTLNDHYDAGL	960	
QY	961	RKLYVNGEFLHNPKEOHLANDLVABABARTLEAKQVPPAKTQCEPVSRRBA	1020	
DB	961	RKLYVNGEFLHNPKEOHLANDLVABABARTLEAKQVPPAKTQCEPVSRRBA	1020	
QY	1021	RAAPDTEPDOSLNLMEAKQALTAETOKSKATKYKSKAVSFDPILLDSLFALBAA	1080	
DB	1021	RAAPDTEPDOSLNLMEAKQALTAETOKSKATKYKSKAVSFDPILLDSLFALBAA	1080	

QY	1081	LEVITAPQOSEKRLAEGEAEKQKQKOLISRYNSALSSETATVNSMLSVQDELRLRY	1140	
DB	1081	LEVITAPQOSEKRLAEGEAEKQKQKOLISRYNSALSSETATVNSMLSVQDELRLRY	1140	
QY	1141	DQASAVWNTNADQKRYTSDAFRAYQOQKTNLFOIGYQKALANRICAIVSHSRSDNTE	1200	
DB	1141	DQASAVWNTNADQKRYTSDAFRAYQOQKTNLFOIGYQKALANRICAIVSHSRSDNTE	1200	
QY	1201	DEQKHAFLTMSSGFQYQWEDLOGVNVTGISASKAERQSKIRKAINGVNAY	1260	
DB	1201	DEQKHAFLTMSSGFQYQWEDLOGVNVTGISASKAERQSKIRKAINGVNAY	1260	
QY	1261	QFRLQGLIQPYGVNRYFIERENYQSEBRYVKTPLAFNRNAGIRVDYFETPDNISV	1320	
DB	1261	QFRLQGLIQPYGVNRYFIERENYQSEBRYVKTPLAFNRNAGIRVDYFETPDNISV	1320	
QY	1321	KPYFVNVYDVSNANVQTTNLTIVQOPFRYQKQKGLKALHPOISAFISQSGSL	1380	
DB	1321	KPYFVNVYDVSNANVQTTNLTIVQOPFRYQKQKGLKALHPOISAFISQSGSL	1380	
QY	1381	GKQOAVGVKLYRW 1394		
DB	1381	GKQOAVGVKLYRW 1394		

RESULT 2
 US-10-645-655-2
 : Sequence 2, Application US/10645655
 : GENERAL INFORMATION:
 : APPLICANT: St. Geme III, Joseph W.
 : TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
 : NUMBER OF SEQUENCES: 9
 : CORRESPONDENCE ADDRESS:
 : ADDRESSES: Flehr, Hobach, Test, Albritton & Herbert
 : STREET: 4 Embarcadero Center, Suite 3400
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States
 : ZIP: 94111-4187
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/10/645,655
 : FILING DATE: 20-Aug-2003
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/296,791
 : FILING DATE: 25-AUG-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Treacatlin, Richard F.
 : REGISTRATION NUMBER: 31,801
 : REFERENCE/DOCKET NUMBER: A-59941/PFT/RMS
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 781-1989
 : TELEFAX: (415) 398-3249
 : TELEX: 910 277999
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1394 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 : US-10-645-655-2
 : Query Match 100.0%; Score 7273; DB 31; Length 1394;
 : Best Local Similarity 100.0%; Pred. No. 0;

Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKCTVRLNPLTACISIGIVSQAMAGHTYFGIDYQYRDPFAENKGFVGAQNIKYNNQ 60
DB 1 MKCTVRLNPLTACISIGIVSQAMAGHTYFGIDYQYRDPFAENKGFVGAQNIKYNNQ 60
QY 61 GQIVGTSMTKAPMIDFVSVRNGVAAVVENQYIVSAHVNGVTDVDFGABANNDDQHRFT 120
DB 61 GQIVGTSMTKAPMIDFVSVRNGVAAVVENQYIVSAHVNGVTDVDFGABANNDDQHRFT 120
QY 121 YKIVKNNYKKNLHPYEDDYNHPRHLKFTVEAAPIDMTSNNGSTYSDRKTYPERVRIG 180
DB 121 YKIVKNNYKKNLHPYEDDYNHPRHLKFTVEAAPIDMTSNNGSTYSDRKTYPERVRIG 180
QY 181 SGQFWRNDQKQDVAGAYHYLTAGNTNORAGANGSYLGADVRKAGEYGLPIAGSK 240
DB 181 SGQFWRNDQKQDVAGAYHYLTAGNTNORAGANGSYLGADVRKAGEYGLPIAGSK 240
QY 241 GDSGSPMFIYDAEKQKWLINGILREGNPBEKENGQVLVRKSYFDSIFPRDLHTSLYTRA 300
DB 241 GDSGSPMFIYDAEKQKWLINGILREGNPBEKENGQVLVRKSYFDSIFPRDLHTSLYTRA 300
QY 301 GNGVYTTISGNDNGQGSITQKSGIPSEIKITLANNSLPLKXCDKYVHNPYDGNISYPRIN 360
DB 301 GNGVYTTISGNDNGQGSITQKSGIPSEIKITLANNSLPLKXCDKYVHNPYDGNISYPRIN 360
QY 361 NGELTYFMQKQSLIFASDINOGAGGLYFBEENFVSHNSQITWOGAGHISYENSITVWK 420
DB 361 NGELTYFMQKQSLIFASDINOGAGGLYFBEENFVSHNSQITWOGAGHISYENSITVWK 420
QY 421 VNGVEHRLSKIGKGLTVQAKGKENGKSVGDGVVLIIEQADQDNKQASEBILGVSR 480
DB 421 VNGVEHRLSKIGKGLTVQAKGKENGKSVGDGVVLIIEQADQDNKQASEBILGVSR 480
QY 481 GTVQANDQKQPTDFEYFEGGGLDLNGHSLTFKRLONTDEGAMVNTNTQAAVYIT 540
DB 481 GTVQANDQKQPTDFEYFEGGGLDLNGHSLTFKRLONTDEGAMVNTNTQAAVYIT 540
QY 541 GNESTIVPNGANNINLKYREIAYNGMFEETDKXNGKRLMYKTYTTEORTLLSGGTV 600
DB 541 GNESTIVPNGANNINLKYREIAYNGMFEETDKXNGKRLMYKTYTTEORTLLSGGTV 600
QY 601 LKGLITCTGKCLFESGRPTPHAVNHLNKRSMBSGIPQGLVMDHDMINKRTPKAEPOIK 660
DB 601 LKGLITCTGKCLFESGRPTPHAVNHLNKRSMBSGIPQGLVMDHDMINKRTPKAEPOIK 660
QY 661 GSAVAVSRNVSIEGNVTVSNNAATFGVVPQOQWCTICTSDMTGLTTCQKVDLDTKVI 720
DB 661 GSAVAVSRNVSIEGNVTVSNNAATFGVVPQOQWCTICTSDMTGLTTCQKVDLDTKVI 720
QY 721 NSIPKQINGSLNTDNATANYKGLAKMANVTLTNHSQPTLSNNAIQIRLSDNSTA 780
DB 721 NSIPKQINGSLNTDNATANYKGLAKMANVTLTNHSQPTLSNNAIQIRLSDNSTA 780
QY 781 TVDNANLNGVHLTSAQPSLKNSHSEQOQDKGTYTLENATIMPBPTTLQNTLIAN 840
DB 781 TVDNANLNGVHLTSAQPSLKNSHSEQOQDKGTYTLENATIMPBPTTLQNTLIAN 840
QY 841 STILINSVASSASNNTPRRSLTETETPTSAHRENTLTVNGKLSQOQTFQTSLSFGYK 900
DB 841 STILINSVASSASNNTPRRSLTETETPTSAHRENTLTVNGKLSQOQTFQTSLSFGYK 900
QY 901 SDKLSNDABEGYILSVNTGKPELTLEDTLIVSKXNOPLSDKLFLENDHVDGAL 960
DB 901 SDKLSNDABEGYILSVNTGKPELTLEDTLIVSKXNOPLSDKLFLENDHVDGAL 960
QY 961 RYKLVKNDGEPRLNPIKEOELHNDLVRAVOAERTLAKOVPEPAKTQGEKXVSRRAA 1020
DB 961 RYKLVKNDGEPRLNPIKEOELHNDLVRAVOAERTLAKOVPEPAKTQGEKXVSRRAA 1020
QY 1021 RAAFPDITLPOSLINALEAKQALTAETOKSRAKTKYKXSKRAVFSDDLDSJALEAA 1080
DB 1021 RAAFPDITLPOSLINALEAKQALTAETOKSRAKTKYKXSKRAVFSDDLDSJALEAA 1080

```

```

QY 1081 LEVIDAPQOSKORLAEBAEKORQKDLISRENSALSELATVNSMLSVQDELDRFLV 1140
DB 1081 LEVIDAPQOSKORLAEBAEKORQKDLISRENSALSELATVNSMLSVQDELDRFLV 1140
QY 1141 DOAQSVAWNTIADKRRYDSCAFRAYOQKNTNRQIGVQKALANGRIGAVFSSRSNDTF 1200
DB 1141 DOAQSVAWNTIADKRRYDSCAFRAYOQKNTNRQIGVQKALANGRIGAVFSSRSNDTF 1200
QY 1201 DEQVYNNATILTMSSGAPOQWGDILQFYNNVGTGISAKSKAEBSQSKIHRAINYGNASY 1260
DB 1201 DEQVYNNATILTMSSGAPOQWGDILQFYNNVGTGISAKSKAEBSQSKIHRAINYGNASY 1260
QY 1261 QRLQGLQIGYRGVNRVFIERENYQSEFERYVTPSLANRNAGIRVDTFTPTNISV 1320
DB 1261 QRLQGLQIGYRGVNRVFIERENYQSEFERYVTPSLANRNAGIRVDTFTPTNISV 1320
QY 1321 KPYFFVNYVDVSNANQVTVNLTVLQOPFGYVQKQEVGKABILFQISAFISKSQSQQL 1380
DB 1321 KPYFFVNYVDVSNANQVTVNLTVLQOPFGYVQKQEVGKABILFQISAFISKSQSQQL 1380
QY 1381 GKQQNVGVKLAGYEM 1394
DB 1381 GKQQNVGVKLAGYEM 1394

```

```

RESULT 3
US-10-687-046-2
; Sequence 2, Application US/10687046
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RET/DCP/DR
; CURRENT APPLICATION NUMBER: US/10/687,046
; PRIOR FILING DATE: 2003-10-15
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-687-046-2

```

Query Match 100.0%; Score 7273; DB 31; Length 1394;
 Best local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKCTVRLNPLTACISIGIVSQAMAGHTYFGIDYQYRDPFAENKGFVGAQNIKYNNQ 60
DB 1 MKCTVRLNPLTACISIGIVSQAMAGHTYFGIDYQYRDPFAENKGFVGAQNIKYNNQ 60
QY 61 GQIVGTSMTKAPMIDFVSVRNGVAAVVENQYIVSAHVNGVTDVDFGABANNDDQHRFT 120
DB 61 GQIVGTSMTKAPMIDFVSVRNGVAAVVENQYIVSAHVNGVTDVDFGABANNDDQHRFT 120
QY 121 YKIVKNNYKKNLHPYEDDYNHPRHLKFTVEAAPIDMTSNNGSTYSDRKTYPERVRIG 180
DB 121 YKIVKNNYKKNLHPYEDDYNHPRHLKFTVEAAPIDMTSNNGSTYSDRKTYPERVRIG 180
QY 181 SGQFWRNDQKQDVAGAYHYLTAGNTNORAGANGSYLGADVRKAGEYGLPIAGSK 240
DB 181 SGQFWRNDQKQDVAGAYHYLTAGNTNORAGANGSYLGADVRKAGEYGLPIAGSK 240
QY 241 GDSGSPMFIYDAEKQKWLINGILREGNPBEKENGQVLVRKSYFDSIFPRDLHTSLYTRA 300
DB 241 GDSGSPMFIYDAEKQKWLINGILREGNPBEKENGQVLVRKSYFDSIFPRDLHTSLYTRA 300

```

QY 301 GNGVTTISGNDGQSSITOKSGIPSEKITLANSPLKXKDYHNPYDGNITSPRLN 360
 DB 301 GNGVTTISGNDGQSSITOKSGIPSEKITLANSPLKXKDYHNPYDGNITSPRLN 360
 QY 361 NGETLYFMDQKQSSIFASDINOGAGLYFEGNFTVPSNSQTMOGAGIHVSNSITVTK 420
 DB 361 NGETLYFMDQKQSSIFASDINOGAGLYFEGNFTVPSNSQTMOGAGIHVSNSITVTK 420
 QY 421 VNGVEHDLRSKIGKGTILHVOAKGENKGSISVGDGVILLBOADDQGNKQAFSEIGVSGR 480
 DB 421 VNGVEHDLRSKIGKGTILHVOAKGENKGSISVGDGVILLBOADDQGNKQAFSEIGVSGR 480
 QY 481 GTVQANDKQPTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMIVHNTTOAANVTIT 540
 DB 481 GTVQANDKQPTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMIVHNTTOAANVTIT 540
 QY 541 GNESTIVLPNGNNINKLDYRKELIANGFGEITDKNGKNGRLMLLYKTTEDRTLLSGGTN 600
 DB 541 GNESTIVLPNGNNINKLDYRKELIANGFGEITDKNGKNGRLMLLYKTTEDRTLLSGGTN 600
 QY 601 LKGDITQTKGKLFFSGRPTPHAYVNLNKRMSSEMGIPQGETVMDHDMINRTFKAEFQIK 660
 DB 601 LKGDITQTKGKLFFSGRPTPHAYVNLNKRMSSEMGIPQGETVMDHDMINRTFKAEFQIK 660
 QY 661 GSAVNSRANSSIEGNFTVSNANATFGVYVNOQKTCITSDPTGLTTCQKXDLTPRKVI 720
 DB 661 GSAVNSRANSSIEGNFTVSNANATFGVYVNOQKTCITSDPTGLTTCQKXDLTPRKVI 720
 QY 721 NSIPKTOJNGSINTCNATAVKGLATNANVTILNHSQFLSNNTQIGNRLSDNSTA 780
 DB 721 NSIPKTOJNGSINTCNATAVKGLATNANVTILNHSQFLSNNTQIGNRLSDNSTA 780
 QY 781 TVDNANINGVHLTDSAQFSLKNSHSEHQIQDCKGTIVTLNNAITWMPSDITLLQNTLKN 840
 DB 781 TVDNANINGVHLTDSAQFSLKNSHSEHQIQDCKGTIVTLNNAITWMPSDITLLQNTLKN 840
 QY 841 STITNSAVSASNNTPRRBSLETETPTSAHRPNLTUNGSLSQCFQGTSLSGYK 900
 DB 841 STITNSAVSASNNTPRRBSLETETPTSAHRPNLTUNGSLSQCFQGTSLSGYK 900
 QY 901 SDKILSNDASGDIYSVRNTGKEPETLEOLTIVSCKNCPLESDKIFLNDHVDAGAL 960
 DB 901 SDKILSNDASGDIYSVRNTGKEPETLEOLTIVSCKNCPLESDKIFLNDHVDAGAL 960
 QY 961 RYKJLVKNDGFEFLNPIKEOELENDLYRABOERTLEKCYEPTANTOTGEPRKYSRRAA 1020
 DB 961 RYKJLVKNDGFEFLNPIKEOELENDLYRABOERTLEKCYEPTANTOTGEPRKYSRRAA 1020
 QY 1021 RAAPFDLTPOSILNATAKQAELETAETOKSAKTKYRSKRAVFSDDLDSQJFALBAA 1080
 DB 1021 RAAPFDLTPOSILNATAKQAELETAETOKSAKTKYRSKRAVFSDDLDSQJFALBAA 1080
 QY 1081 LEVIDAPQOSEKDRLOAEAEKQKQKDLISRYNSALSLSATVSNMLSVODELRIFFV 1140
 DB 1081 LEVIDAPQOSEKDRLOAEAEKQKQKDLISRYNSALSLSATVSNMLSVODELRIFFV 1140
 QY 1141 DOAGSAVNTIAODKRRDSDAPRAYOQKTNLROIGYQALANGKIGAVSHSHSNTPT 1200
 DB 1141 DOAGSAVNTIAODKRRDSDAPRAYOQKTNLROIGYQALANGKIGAVSHSHSNTPT 1200
 QY 1201 DEOVKNATILTMASGFAOYQMSDQFQGVNIGCISASMAEGRKIKHAKIANYVANSY 1260
 DB 1201 DEOVKNATILTMASGFAOYQMSDQFQGVNIGCISASMAEGRKIKHAKIANYVANSY 1260
 QY 1261 QFRIGQIGIQPYFGVNRYPILRENYQSEEVRYKTPSLAFNRKYNAGRVDYTFPTDINISV 1320
 DB 1261 QFRIGQIGIQPYFGVNRYPILRENYQSEEVRYKTPSLAFNRKYNAGRVDYTFPTDINISV 1320
 QY 1321 KPYFRANVUVSNAMQTVNLTVLOQPFERYWQKVGKLELHFOJISAFISQSGSOL 1380
 DB 1321 KPYFRANVUVSNAMQTVNLTVLOQPFERYWQKVGKLELHFOJISAFISQSGSOL 1380
 QY 1381 GKQONVGVKLYGEM 1394
 DB 1381 GKQONVGVKLYGEM 1394

DB 1381 GKQONVGVKLYGEM 1394
 RESULT 4
 PCT-US03-05226-7
 ; Sequence 7, Application EC/TUS0305226
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RT/DCF/DBR
 ; CURRENT APPLICATION NUMBER: PCT/US03/05226
 ; CURRENT FILING DATE: 2003-02-18
 ; PRIOR APPLICATION NUMBER: US 08/296,731
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 1395
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; PCT-US03-05226-7
 Query Match 99.6%; Score 7246.5; DB 1; Length 1395;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1393; Conservative 0; Mismatches 0; Indels 3; Gaps 2;
 QY 1 MKKTVFRLNPLFACISIGIVSOAMAGHTYFGIDYQYRDFAEKNGKFTVQAKNIKYNNQ 60
 DB 1 MKKTVFRLNPLFACISIGIVSOAMAGHTYFGIDYQYRDFAEKNGKFTVQAKNIKYNNQ 60
 QY 61 GQIVGSMTKAPMIDSEVSRNGVAALVENQYIVSAHVNGYTDVDFAGAGNPPDOHRT 120
 DB 61 GQIVGSMTKAPMIDSEVSRNGVAALVENQYIVSAHVNGYTDVDFAGAGNPPDOHRT 120
 QY 121 YKIVKNNYKKNQNHPEDDYENPLHFYTEAPIDMTSNNGSYSDRKYPERVRIG 180
 DB 121 YKIVKNNYKKNQNHPEDDYENPLHFYTEAPIDMTSNNGSYSDRKYPERVRIG 180
 QY 181 SGROFMRNDQDQVAGAYHVLTAGNHNDRQAGNGSYLGGDYRAXGEYGLPIAGSK 240
 DB 181 SGROFMRNDQDQVAGAYHVLTAGNHNDRQAGNGSYLGGDYRAXGEYGLPIAGSK 240
 QY 241 GDSGSMFTIYAEKQKMLINGILREGNPEKENGQVYKSYFPELPERDLTSLYTA 300
 DB 241 GDSGSMFTIYAEKQKMLINGILREGNPEKENGQVYKSYFPELPERDLTSLYTA 300
 QY 301 GDSGSMFTIYAEKQKMLINGILREGNPEKENGQVYKSYFPELPERDLTSLYTA 300
 DB 301 GDSGSMFTIYAEKQKMLINGILREGNPEKENGQVYKSYFPELPERDLTSLYTA 300
 QY 301 GNGVTTISGNDGQSSITOKSGIPSEKITLANSPLKXKDYHNPYDGNITSPRLN 360
 DB 301 GNGVTTISGNDGQSSITOKSGIPSEKITLANSPLKXKDYHNPYDGNITSPRLN 360
 QY 361 NGETLYFMDQKQSSIFASDINOGAGLYFEGNFTVPSNSQTMOGAGIHVSNSITVTK 420
 DB 361 NGETLYFMDQKQSSIFASDINOGAGLYFEGNFTVPSNSQTMOGAGIHVSNSITVTK 420
 QY 421 VNGVEHDLRSKIGKGTILHVOAKGENKGSISVGDGVILLBOADDQGNKQAFSEIGVSGR 480
 DB 421 VNGVEHDLRSKIGKGTILHVOAKGENKGSISVGDGVILLBOADDQGNKQAFSEIGVSGR 480
 QY 481 GTVQANDKQPTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMIVHNTTOAANVTIT 540
 DB 481 GTVQANDKQPTDKFYFGRGRLDNGHSLTFKRIQNTDEGAMIVHNTTOAANVTIT 540
 QY 541 GNESTIVLPNGNNINKLDYRKELIANGFGEITDKNGKNGRLMLLYKTTEDRTLLSGGTN 600
 DB 541 GNESTIVLPNGNNINKLDYRKELIANGFGEITDKNGKNGRLMLLYKTTEDRTLLSGGTN 600
 QY 601 LKGDITQTKGKLFFSGRPTPHAYVNLNKRMSSEMGIPQGETVMDHDMINRTFKAEFQIK 660
 DB 601 LKGDITQTKGKLFFSGRPTPHAYVNLNKRMSSEMGIPQGETVMDHDMINRTFKAEFQIK 660

```

QY 661 GGSAYVSRNVSIEGNTVTUSNNANATFCVNPNOQNTITCRSDWGTITTCOKVLDITDKYI 720
DB 661 GGSAYVSRNVSIEGNTVTUSNNANATFCVNPNOQNTITCRSDWGTITTCOKVLDITDKYI 720
QY 721 NSIPKTOINGSINTLNATANYKGLAKNGVNTLTNSOFTLSNNATOIGIRLSNSTA 780
DB 721 NSIPKTOINGSINTLNATANYKGLAKNGVNTLTNSOFTLSNNATOIGIRLSNSTA 780
QY 781 TYDNANLNGVHLTDSAQFSLKNHSHFHOIQDGKTTVTLLENATWMSDTTLQNTLTNN 840
DB 781 TYDNANLNGVHLTDSAQFSLKNHSHFHOIQDGKTTVTLLENATWMSDTTLQNTLTNN 840
QY 841 STTTLSAIVASASNNTP--RRSLTETTPPSAHRNTLTLYNGLSGQGFQFTSLG 898
DB 841 STTTLSAIVASASNNTP--RRSLTETTPPSAHRNTLTLYNGLSGQGFQFTSLG 898
QY 899 YKSDKLKLSNDAEGDYLISVRNTGKEPETTLTQTLVSKONQPLSDKLKFTLENDHVDAG 958
DB 899 YKSDKLKLSNDAEGDYLISVRNTGKEPETTLTQTLVSKONQPLSDKLKFTLENDHVDAG 958
QY 901 YKSDKLKLSNDAEGDYLISVRNTGKEPETTLTQTLVSKONQPLSDKLKFTLENDHVDAG 960
DB 901 YKSDKLKLSNDAEGDYLISVRNTGKEPETTLTQTLVSKONQPLSDKLKFTLENDHVDAG 960
QY 959 ALRYKLKVNDEGFRHNPRIKCEJELNDLVRAEQERTLEAKZYEPFAATQGTGEPKYSRR 1018
DB 959 ALRYKLKVNDEGFRHNPRIKCEJELNDLVRAEQERTLEAKZYEPFAATQGTGEPKYSRR 1018
QY 961 ALRYKLKVNDEGFRHNPRIKCEJELNDLVRAEQERTLEAKZYEPFAATQGTGEPKYSRR 1026
DB 961 ALRYKLKVNDEGFRHNPRIKCEJELNDLVRAEQERTLEAKZYEPFAATQGTGEPKYSRR 1026
QY 1019 AABAAPPTLPDQSLNALAKOAEITATQSKAKTKVRSRAVSDPILQSLFALE 1078
DB 1019 AABAAPPTLPDQSLNALAKOAEITATQSKAKTKVRSRAVSDPILQSLFALE 1078
QY 1021 AABAAPPTLPDQSLNALAKOAEITATQSKAKTKVRSRAVSDPILQSLFALE 1086
DB 1021 AABAAPPTLPDQSLNALAKOAEITATQSKAKTKVRSRAVSDPILQSLFALE 1086
QY 1079 AALEVIDAPQSEKORLQOEAERKQKQKDIISRYNSALSSELSATVMSLSTVODELRL 1138
DB 1079 AALEVIDAPQSEKORLQOEAERKQKQKDIISRYNSALSSELSATVMSLSTVODELRL 1138
QY 1139 FVDQASAVMTNIAQDRRYSDAPFAYOQKTNLQIGVQALANGRIGAVSSEBSJN 1198
DB 1139 FVDQASAVMTNIAQDRRYSDAPFAYOQKTNLQIGVQALANGRIGAVSSEBSJN 1198
QY 1141 FVDQASAVMTNIAQDRRYSDAPFAYOQKTNLQIGVQALANGRIGAVSSEBSJN 1199
DB 1141 FVDQASAVMTNIAQDRRYSDAPFAYOQKTNLQIGVQALANGRIGAVSSEBSJN 1199
QY 1199 TPEDQVKNATLTWMSGPAQYQCGDIQFGVNWGJISASQMEESQRIKHKANLYGVA 1258
DB 1199 TPEDQVKNATLTWMSGPAQYQCGDIQFGVNWGJISASQMEESQRIKHKANLYGVA 1258
QY 1259 SYQFRLGQIGIPIYGVNRYEIRENYQSEEVRYKTPSLAENRYNAGIRVDYTPPTDNI 1318
DB 1259 SYQFRLGQIGIPIYGVNRYEIRENYQSEEVRYKTPSLAENRYNAGIRVDYTPPTDNI 1318
QY 1260 SYQFRLGQIGIPIYGVNRYEIRENYQSEEVRYKTPSLAENRYNAGIRVDYTPPTDNI 1319
DB 1260 SYQFRLGQIGIPIYGVNRYEIRENYQSEEVRYKTPSLAENRYNAGIRVDYTPPTDNI 1319
QY 1319 SVKPYFFVNYVDVSNANVQTTNLTVLOQPFGRYQKEVSLKAEILHFOISAFISQSGS 1378
DB 1319 SVKPYFFVNYVDVSNANVQTTNLTVLOQPFGRYQKEVSLKAEILHFOISAFISQSGS 1378
QY 1379 QLGKQONGVKLGKYM 1394
DB 1379 QLGKQONGVKLGKYM 1394
QY 1380 QLGKQONGVKLGKYM 1395
DB 1380 QLGKQONGVKLGKYM 1395

```

RESULT 5

```

US-10-687-046-7
Sequence 7, Application US/10687046
GENERAL INFORMATION:
APPLICANT: St. Geme, Joseph W.
TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
FILE REFERENCE: A-59941-1/RPT/DOF/DHR
CURRENT APPLICATION NUMBER: US/10/687,046
CURRENT FILING DATE: 2003-10-15
PRIOR APPLICATION NUMBER: US/10/080,505
PRIOR FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: US 08/296,791
PRIOR FILING DATE: 1994-10-25
PRIOR APPLICATION NUMBER: US 09/839,996
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 1395
TYPE: PRT

```

ORGANISM: Haemophilus influenzae
US-10-687-046-7

Query Match 99.6%; Score 7246.5; DB 31; Length 1395;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

```

QY 1 MKTYFRANLTLCTSLGIVQAAAGTTTGTIDYQYDFAFMKGKTTVAQNTIVYVKQ 60
DB 1 MKTYFRANLTLCTSLGIVQAAAGTTTGTIDYQYDFAFMKGKTTVAQNTIVYVKQ 60
QY 61 GOLVGTSMKAPMIDSVYSRNGVAALVENQYIVSAHVNGTVDVDAEKNPDQHRFT 120
DB 61 GOLVGTSMKAPMIDSVYSRNGVAALVENQYIVSAHVNGTVDVDAEKNPDQHRFT 120
QY 121 YKIVRNNYKQDNLHAYEDDYHNPRLKFTYEAAPIDMTSNMGSTYSDETKYPERVRIG 180
DB 121 YKIVRNNYKQDNLHAYEDDYHNPRLKFTYEAAPIDMTSNMGSTYSDETKYPERVRIG 180
QY 181 SCRGPFMRNDQKGDQYAGATHTLTNGTTHQKQAGNGSYLGSVRRAGYGPPLIAGSK 240
DB 181 SCRGPFMRNDQKGDQYAGATHTLTNGTTHQKQAGNGSYLGSVRRAGYGPPLIAGSK 240
QY 241 GDSGPMETIDAEKQKMLNGLIREGNPFEGKNGFOLVRKSYDELFERDLHTSLYTRA 300
DB 241 GDSGPMETIDAEKQKMLNGLIREGNPFEGKNGFOLVRKSYDELFERDLHTSLYTRA 300
QY 301 GNGVYTTISGNDNGQSSITQKSGIDSEIKITLAMSLLPKXQKQKHNRYDGENYISPRLN 360
DB 301 GNGVYTTISGNDNGQSSITQKSGIDSEIKITLAMSLLPKXQKQKHNRYDGENYISPRLN 360
QY 361 NERTLYEPNDQKQSLFASDINQAGAGLYPENGFTVSPNSQKQAGHVSNSVYWK 420
DB 361 NERTLYEPNDQKQSLFASDINQAGAGLYPENGFTVSPNSQKQAGHVSNSVYWK 420
QY 421 VNGVZHEDELKTIKGTATVQAKGNGKGSISVGDQKYLEQOADDQKQKQFSEIGVSGR 480
DB 421 VNGVZHEDELKTIKGTATVQAKGNGKGSISVGDQKYLEQOADDQKQKQFSEIGVSGR 480
QY 481 GTVQALNDKQPDITDKYFGFGRGRDLNAGSLTFKRIQNTDEGAMLYNHTTQAAVYTT 540
DB 481 GTVQALNDKQPDITDKYFGFGRGRDLNAGSLTFKRIQNTDEGAMLYNHTTQAAVYTT 540
QY 541 GNESTIVLPNANNINKLDYKXELAVNGMFGTDQKNGKGLNLYPTTEDRTLLSGGN 600
DB 541 GNESTIVLPNANNINKLDYKXELAVNGMFGTDQKNGKGLNLYPTTEDRTLLSGGN 600
QY 601 LKGDITQTKGKLFFSGRPPTPHAVNHLNKRWESEMEGIDQELIYWDHWINRTPFAENFOIX 660
DB 601 LKGDITQTKGKLFFSGRPPTPHAVNHLNKRWESEMEGIDQELIYWDHWINRTPFAENFOIX 660
QY 661 GGSAYVSRNVSIEGNTVTUSNNANATFCVNPNOQNTITCRSDWGTITTCOKVLDITDKYI 720
DB 661 GGSAYVSRNVSIEGNTVTUSNNANATFCVNPNOQNTITCRSDWGTITTCOKVLDITDKYI 720
QY 721 NSIPKTOINGSINTLNATANYKGLAKNGVNTLTNSOFTLSNNATOIGIRLSNSTA 780
DB 721 NSIPKTOINGSINTLNATANYKGLAKNGVNTLTNSOFTLSNNATOIGIRLSNSTA 780
QY 781 TYDNANLNGVHLTDSAQFSLKNHSHFHOIQDGKTTVTLLENATWMSDTTLQNTLTNN 840
DB 781 TYDNANLNGVHLTDSAQFSLKNHSHFHOIQDGKTTVTLLENATWMSDTTLQNTLTNN 840
QY 841 STTTLSAIVASASNNTP--RRSLTETTPPSAHRNTLTLYNGLSGQGFQFTSLG 898
DB 841 STTTLSAIVASASNNTP--RRSLTETTPPSAHRNTLTLYNGLSGQGFQFTSLG 898
QY 899 YKSDKLKLSNDAEGDYLISVRNTGKEPETTLTQTLVSKONQPLSDKLKFTLENDHVDAG 958
DB 899 YKSDKLKLSNDAEGDYLISVRNTGKEPETTLTQTLVSKONQPLSDKLKFTLENDHVDAG 958
QY 901 ALRYKLKVNDEGFRHNPRIKCEJELNDLVRAEQERTLEAKZYEPFAATQGTGEPKYSRR 1018
DB 901 ALRYKLKVNDEGFRHNPRIKCEJELNDLVRAEQERTLEAKZYEPFAATQGTGEPKYSRR 1018

```

Db 96: ALRYKLVKNDGERFLNPLIKQELHNDLVRABQAEFTLEAKQVEPTAKTQTEGPKKISR 1020
 Qy 1039 AARAFEDTLIPDQSLINLLEAKQELTLETOKSKAKTKKVSRAVEFSDP.LDDGFLALE 1078
 Db 1032: AARAFEDTLIPDQSLINLLEAKQELTLETOKSKAKTKKVSRAVEFSDP.LDDGFLALE 1080
 Qy 1079 AALEYITAPQOSEKCKLAGEAEKQKQKDLISRYNSALSELATYNSMLSVQDELDEL 1138
 Db 1081: AALEYIDAPQOSEKCKLAGEAEKQKQKDLISRYNSALSELATYNSMLSVQDELDEL 1140
 Qy 1139 FVDAQASAVWCTNADQKRYDSDAFRAYQQQKTNLRQIGVQKALANGRIGAVFSSRSDN 1198
 Db 1141: FVDAQASAVWCTNADQKRYDSDAFRAYQQQKTNLRQIGVQKALANGRIGAVFSSRSDN 1199
 Qy 1199 FDEQVYKNAHLLTMSGFAQYQWMDLQPGVNVGTGSIASXMAEBSRKIHRRALNYGMA 1258
 Db 1200 FDEQVYKNAHLLTMSGFAQYQWMDLQPGVNVGTGSIASXMAEBSRKIHRRALNYGMA 1259
 Qy 1259 SVQRLGGLGIPRYQVRYFIERENYQSEEVRYKTPSLAFNRINAGISVDYTPPTDNI 1318
 Db 1260 SVQRLGGLGIPRYQVRYFIERENYQSEEVRYKTPSLAFNRINAGISVDYTPPTDNI 1319
 Qy 1319 SVKEFFVNVYDVSNANVQTTVNLTVLQCPFGRYQKQKVLKAEILHFOISAFISKQGS 1378
 Db 1320 SVKEFFVNVYDVSNANVQTTVNLTVLQCPFGRYQKQKVLKAEILHFOISAFISKQGS 1379
 Qy 1379 QLGKQNVGVKLGKYM 1394
 Db 1380 QLGKQNVGVKLGKYM 1395

RESULT 6 PCT-US03-05226-9

; Sequence 9, Application PC/TUS0305226
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/AFI/DCF/DHR
 ; CURRENT APPLICATION NUMBER: PCT/US03/05226
 ; CURRENT FILING DATE: 2003-02-18
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 9
 ; LENGTH: 1434
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (568)..(568)
 ; OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
 ; NAME/KEY: misc_feature
 ; LOCATION: (1702)..(1702)
 ; OTHER INFORMATION: "n" at position 1702 can be any base.
 PCT-US03-05226-9

Query Match 84.1%; Score 6115; DB 1; Length 434;
 Best Local Similarity 82.9%; Pred. No. 0;
 Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 10;

Qy 1 MKKTVFRLNPLIACISLIGVQAMAGHTFGIDYQYRDPFAEANKGKFTVGAONIKVYKQ 60
 Db 1 MKKTVFRLNPLIACISLIGVQAMAGHTFGIDYQYRDPFAEANKGKFTVGAONIKVYK 60
 Qy 61 GGLVSTSTTAAKIMIDSSVSRNGAALYENQYVSAHNQVCTYVPGAGNNPOHFT 120
 Db 61 GGLVSTSTTAAKIMIDSSVSRNGAALYENQYVSAHNQVCTYVPGAGNNPOHFT 120
 Qy 121 YKIVKNNYKKNMIAHYEDDYHNPRLHKEFTEAAPIDMTSNMNSITYSRITXYPERRVIG 180

Db 121 YKIVKNNYKKNMIAHYEDDYHNPRLHKEFTEAAPIDMTSNMNSITYSRITXYPERRVIG 180
 Qy 181 SGRQFRRNDQKDPQVAGAYHYLTAGNTHQVQAGNGSYSLGQVYKXAGEXGLPIAGSK 240
 Db 181 SGRQVYRDDQNNRTYLSIDGYNTLTGANTYNSGSGGQGYVYVAGDIRKXGQVGLPIASSF 240
 Qy 241 GDSGSPMTIYAEKQKMLINGILRGNPFGKENGQVLYKSYFPELPERDLTSLYTA 300
 Db 241 GDSGSPMTIYAEKQKMLINGILRGNPFGKENGQVLYKSYFPELPERDLTSLYTA 300
 Qy 301 GNGVYITSGNMGKSGSTQKSGIPSEIKITLNMMSLEAKKQKYNHPYDGPNTYSPRLN 360
 Db 301 GNGVYITSGNMGKSGSTQKSGIPSEIKITLNMMSLEAKKQKYNHPYDGPNTYSPRLN 360
 Qy 361 NSELTLYMDQKQSLIFASDINQAGAGLYPEGNFTVSPNSNQTWQAGIHYSENSYTWK 420
 Db 359 NKNITFYFQEBLIGSITLTTDIDQAGAGLYPEGNFTVSPNSNQTWQAGIHYSENSYTWK 418
 Qy 421 VNGTEHRLSKTIGKTLHYQAKENGKSI SVBGKVTILEQQADDDQKQAFSEIGLYSR 480
 Db 419 VNGTEHRLSKTIGKTLHYQAKENGKSI SVBGKVTILEQQADDDQKQAFSEIGLYSR 478
 Qy 481 GTVQANDKQPDTPKPYGFGGRLTLAKESLAPKRIQNTDEGAMIYVNHNTQAAVYTT 540
 Db 479 GTVQANDKQPDTPKPYGFGGRLTLAKESLAPKRIQNTDEGAMIYVNHNTQAAVYTT 538
 Qy 541 GNESTIVLPNGNNINKLDYRKELIAYNGFGETDKNKNKELMILYKPTTEDRTLLSGGT 599
 Db 539 GNESTIVLPNGNNINKLDYRKELIAYNGFGETDKNKNKELMILYKPTTEDRTLLSGGT 598
 Qy 600 NLKQDITQTKGLFSSRPPTPRAYNHLNKRSEMEGIPQGLVWMDIMNTPRASNFOI 659
 Db 599 NLKQDITQTKGLFSSRPPTPRAYNHLNKRSEMEGIPQGLVWMDIMNTPRASNFOI 658
 Qy 660 KGSAAVSNVSSILEGNTVSNANATPGVYVQONTICRSDDWGLTTCQKVDLTDYV 719
 Db 659 KGSAAVSNVSSILEGNTVSNANATPGVYVQONTICRSDDWGLTTCQKVDLTDYV 718
 Qy 720 INSIPATQINSINLTQNTATNAYKGLAKANGVTLNHSOFTLSNNAFOINRILS 775
 Db 719 INSIPATQINSINLTQNTATNAYKGLAKANGVTLNHSOFTLSNNAFOINRILS 778
 Qy 776 ---DNST-----ATVDNANLGNVHLTDSKPS 800
 Db 779 ARVNNATLMGCVNLADTSRFTLSNQAQIGTISLHQDQATVDNANLGNVHLTDSKPS 838
 Qy 801 LKNSHSHQIQDQKGTITLENATWTWESDITLQNTLNNSTITLNSAYSAASNTP-RR 859
 Db 839 LKNSHSHQIQDQKGTITLENATWTWESDITLQNTLNNSTITLNSAYSAASNTP-RR 898
 Qy 860 RSLFETITPTSAERFNTLTVNGKLSGGTFOFTSSIFGYKSDKLSNDEGGYTLTYR 919
 Db 899 RSLFETITPTSAERFNTLTVNGKLSGGTFOFTSSIFGYKSDKLSNDEGGYTLTYR 958
 Qy 920 NTGKEPTELEQTLVESKQNPISDKLKEFTLENDHVAQALRYKLVKNDGFFRLHNPKE 979
 Db 959 NTGKEPTELEQTLVESKQNPISDKLKEFTLENDHVAQALRYKLVKNDGFFRLHNPKE 1018
 Qy 980 QELHNDLVRABQAEFTLEAKQVEPTAKTQTEGPKKISR 1039
 Db 1019 QELHNDLVRABQAEFTLEAKQVEPTAKTQTEGPKKISR 1075
 Qy 1040 KQAELETAETQSKAKTKKVSRAVEFSDP.LDDGFLALE 1095
 Db 1076 KQAELETAETQSKAKTKKVSRAVEFSDP.LDDGFLALE 1135
 Qy 1096 AQ-EEAEKQKQKDLISRYNSALSELATYNSMLSVQDELDELFVDAQASAVWCTNAD 1154
 Db 1136 AQEEAEKQKQKDLISRYNSALSELATYNSMLSVQDELDELFVDAQASAVWCTNAD 1195
 Qy 1155 KRYSDAFRAYQQQKTNLRQIGVQKALANGRIGAVFSSRSDNFTFDEQVYKNAHLLTMS 1214

Db 1196 KRRYSDAFRAY-QQKTNLRQIGVQKALDNGRIGAVFSHSRSDNTFDEQVKNATLTWMS 1254
 Qy 1215 GFAQYQWMDLQFGVNVGTGISASRNAEQQSKHRRALNYGVNASYQFRLGQIGIOPYG 1274
 Db 1255 GFAQYQWMDLQFGVNVGTGISASRNAEQQSKHRRALNYGVNASYQFRLGQIGIOPYG 1314
 Qy 1275 VNRYPFERENYQSEEVAVKTPSLAFNRYNAGIRVDYTFPTDNIISVKEYPFVNVVNSA 1334
 Db 1315 VNRYPFERENYQSEEVAVKTPSLAFNRYNAGIRVDYTFPTDNIISVKEYPFVNVVNSA 1374
 Qy 1335 NVQTTVNTLTVLQOPFGVYQKQVGVKAEILHFOISAFISKSQSQSLQKQNVGKLGYN 1394
 Db 1375 NVQTTVNTLTVLQOPFGVYQKQVGVKAEILHFOISAFISKSQSQSLQKQNVGKLGYN 1434

RESULT 7
 US-10-687-046-9
 Sequence 9, Application US/10687046
 GENERAL INFORMATION:
 APPLICANT: St. Geme, Joseph M.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RFT/DCF/DHR
 CURRENT APPLICATION NUMBER: US/10/687,046
 CURRENT FILING DATE: 2003-10-15
 PRIOR APPLICATION NUMBER: US/10/080,505
 PRIOR FILING DATE: 2002-02-22
 PRIOR APPLICATION NUMBER: US 08/296,791
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: Patent version 3.1
 SEQ ID NO 9
 LENGTH: 1434
 TYPE: PRT
 ORGANISM: Haemophilus influenzae
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (568)..(568)
 OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
 NAME/KEY: misc_feature
 LOCATION: (1702)..(1702)
 OTHER INFORMATION: "n" at position 1702 can be any base.
 US-10-687-046-9

Query Match 84.1%; Score 6115; DB 31; Length 1434;
 Best Local Similarity 82.9%; Pred. No. 0;
 Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 10;

Qy 1 MKTVFLNFTLACISIGIVSQAAGHTYFGIDYQYRDPRAENKKGFTYGAQNTIKYVNG 60
 Db 1 MKTVFLNFTLACISIGIVSQAAGHTYFGIDYQYRDPRAENKKGFTYGAQNTIKYVNG 60
 Qy 61 GQVGTSMTRKAPMIDFSVSRKGVYALVGGQITVSAHNGYGVNDVFGAEQNDQRF 120
 Db 61 GQVGTSMTRKAPMIDFSVSRKGVYALVGGQITVSAHNGYGVNDVFGAEQNDQRF 120
 Qy 121 YKIVKNNYKKNLHPYEDDYHNPRLHKEVTEAPIDMTSNNGSTYSRTKTYPERVLG 180
 Db 121 YKIVKNNYKKNLHPYEDDYHNPRLHKEVTEAPIDMTSNNGSTYSRTKTYPERVLG 180
 Qy 181 SGQFWENDQKGDVAGAYHYLTAGNTNORGANGKSYLGDVAKAGVGPPLNGSK 240
 Db 181 SGQFWENDQKGDVAGAYHYLTAGNTNORGANGKSYLGDVAKAGVGPPLNGSK 240
 Qy 241 GDSGPMETLDAEKQKLLINGILBENPREGNGFQVLYKSYFDEIFERDLHSLTBA 300
 Db 241 GDSGPMETLDAEKQKLLINGILBENPREGNGFQVLYKSYFDEIFERDLHSLTBA 300
 Qy 301 GNGVYTTISGANDQGSITQKSGIPSEIKITLANNSLPLKEDKRVANPRYDGNITSPRLN 360
 Db 301 GNGVYTTISGANDQGSITQKSGIPSEIKITLANNSLPLKEDKRVANPRYDGNITSPRLN 360

Qy 361 NGETLYPMQKQSLIPASDINQAGLYPEGNFTVSPNSQNMQAGIHVSENSTYTWK 420
 Db 361 NGETLYPMQKQSLIPASDINQAGLYPEGNFTVSPNSQNMQAGIHVSENSTYTWK 420
 Qy 421 VNGVEHRLSKIGKGLHVAQKNGKSGISVGDGVVILEQADQGNKQAFSEIGLVSGR 480
 Db 421 VNGVEHRLSKIGKGLHVAQKNGKSGISVGDGVVILEQADQGNKQAFSEIGLVSGR 480
 Qy 481 GTVQLNDQKQPTDKEFPFGKRLDNGHSLTFKRIQNTDEGAMVNHNTQAAVNTIT 540
 Db 481 GTVQLNDQKQPTDKEFPFGKRLDNGHSLTFKRIQNTDEGAMVNHNTQAAVNTIT 540
 Qy 541 GNESTVLP-NGNNYKLDYRKEKIAVNGAFEGDQKKNHGRANLYPTEDEETLLSGGT 599
 Db 541 GNESTVLP-NGNNYKLDYRKEKIAVNGAFEGDQKKNHGRANLYPTEDEETLLSGGT 599
 Qy 599 NIKGDIQTQTKLFFESGRPTPHAVNHLNKKSEMEGIPOGEIWDHDMINRPFKAENFOI 659
 Db 599 NIKGDIQTQTKLFFESGRPTPHAVNHLNKKSEMEGIPOGEIWDHDMINRPFKAENFOI 659
 Qy 660 KQGSAYVSRVSSIEGNTVSNANATFVGVNQNTICTSDMTGLTQKAYDLTDYK 719
 Db 660 KQGSAYVSRVSSIEGNTVSNANATFVGVNQNTICTSDMTGLTQKAYDLTDYK 719
 Qy 719 IDSLPTQJNGSINLNDNATVINGLAKMGNVTLNHSQTLSSNATQIGNIKSNHAN 778
 Db 719 IDSLPTQJNGSINLNDNATVINGLAKMGNVTLNHSQTLSSNATQIGNIKSNHAN 778
 Qy 776 ---DNST-----ATVDNANLNGVHLTSDAQS 800
 Db 776 ---DNST-----ATVDNANLNGVHLTSDAQS 800
 Qy 801 LNSHFSHQIQSGKQTTVLENATWYPSDTLLONTLANSITLNSAYASANNTP-RR 859
 Db 801 LNSHFSHQIQSGKQTTVLENATWYPSDTLLONTLANSITLNSAYASANNTP-RR 859
 Qy 859 RSLFETTPYSAHRNNTLVNGKLSQGGTPTTPTLPTFYSQDKLSNABEDYTLSTR 919
 Db 859 RSLFETTPYSAHRNNTLVNGKLSQGGTPTTPTLPTFYSQDKLSNABEDYTLSTR 919
 Qy 919 NGKEPETLBOULTLVESKNOPLSDPKLFTLNDHVDAGALRYKLVNDSEFLNHPKE 979
 Db 919 NGKEPETLBOULTLVESKNOPLSDPKLFTLNDHVDAGALRYKLVNDSEFLNHPKE 979
 Qy 980 QELHNDIVRAQABRLTBAKQVEPTATOTGEKRVSRRAAAAPPTLTDQSLMALEA 1039
 Db 980 QELHNDIVRAQABRLTBAKQVEPTATOTGEKRVSRRAAAAPPTLTDQSLMALEA 1039
 Qy 1039 QELHNDIVRAQABRLTBAKQVEPTATOTGEKRVSRRAAAAPPTLTDQSLMALEA 1075
 Db 1039 QELHNDIVRAQABRLTBAKQVEPTATOTGEKRVSRRAAAAPPTLTDQSLMALEA 1075
 Qy 1040 KQAEITLQKRAKTKKVRKRAV---FSDPLDQS-LFALEALVIVAPQSSKRL 1095
 Db 1040 KQAEITLQKRAKTKKVRKRAV---FSDPLDQS-LFALEALVIVAPQSSKRL 1095
 Qy 1095 KQAEITLQKRAKTKKVRKRAV---FSDPLDQS-LFALEALVIVAPQSSKRL 1154
 Db 1095 KQAEITLQKRAKTKKVRKRAV---FSDPLDQS-LFALEALVIVAPQSSKRL 1154
 Qy 1155 KRRYSDAFRAY-QQKTNLRQIGVQKALDNGRIGAVFSHSRSDNTFDEQVKNATLTWMS 1254
 Db 1155 KRRYSDAFRAY-QQKTNLRQIGVQKALDNGRIGAVFSHSRSDNTFDEQVKNATLTWMS 1254
 Qy 1255 GFAQYQWMDLQFGVNVGTGISASRNAEQQSKHRRALNYGVNASYQFRLGQIGIOPYG 1314
 Db 1255 GFAQYQWMDLQFGVNVGTGISASRNAEQQSKHRRALNYGVNASYQFRLGQIGIOPYG 1314
 Qy 1315 VNRYPFERENYQSEEVAVKTPSLAFNRYNAGIRVDYTFPTDNIISVKEYPFVNVVNSA 1374
 Db 1315 VNRYPFERENYQSEEVAVKTPSLAFNRYNAGIRVDYTFPTDNIISVKEYPFVNVVNSA 1374
 Qy 1375 NVQTTVNTLTVLQOPFGVYQKQVGVKAEILHFOISAFISKSQSQSLQKQNVGKLGYN 1434
 Db 1375 NVQTTVNTLTVLQOPFGVYQKQVGVKAEILHFOISAFISKSQSQSLQKQNVGKLGYN 1434

```

RESULT 8
PCT-US03-05226-17
; Sequence 17, Application PC/TUS0305226
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: PCT/US03/05226
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
PCT-US03-05226-17

```

Query Match 80.1%; Score 5827.5; DB 1; Length 1411;

Best Local Similarity 79.5%; Ered No. 0;

Matches 1133; Conservative 107; Mismatches 139; Indels 47; Gaps 12;

```

QY 1 MKTVPRLNLTACISLIGVSCAWAGHTYFGIDYQYRDEAKKGFVGAONIKVTKQ 60
DB 1 MKTVPRLNLTACISLIGVSCAWAGHTYFGIDYQYRDEAKKGFVGAONIKVTKN 60
QY 61 GVLGTSMTAPMIDPFSVSRNGVALYENQYIVANVGYTDVDFGAGNPNPDRFT 120
DB 61 GVLGTSMTAPMIDPFSVSRNGVALYENQYIVANVGYTDVDFGAGNPNPDRFT 120
QY 121 YKIVKNNYKQCKAPYEDDYHNPRLHKFVTEAEIDMTSMNKSTYSRDTKCPERYA 180
DB 121 YKIVKNNYKQCKAPYEDDYHNPRLHKFVTEAEIDMTSMNKSTYSRDTKCPERYA 180
QY 122 YKIVKNNYKQCKAPYEDDYHNPRLHKFVTEAEIDMTSMNKSTYSRDTKCPERYA 180
DB 122 YKIVKNNYKQCKAPYEDDYHNPRLHKFVTEAEIDMTSMNKSTYSRDTKCPERYA 180
QY 181 SGQFWRNDQNDQVAGAYHYLTAGNTNORAGAGNSYVSGDVRKAGEYEPFLIAGSK 240
DB 181 SGQFWRNDQNDQVAGAYHYLTAGNTNORAGAGNSYVSGDVRKAGEYEPFLIAGSK 240
QY 241 SPGSPMEITDAEKQWILNGLIEGNPEFGKNGFOLVRSYSDHIFERDLHSLVTRA 300
DB 241 SPGSPMEITDAEKQWILNGLIEGNPEFGKNGFOLVRSYSDHIFERDLHSLVTRA 300
QY 241 SPGSPMEITDAEKQWILNGLIEGNPEFGKNGFOLVRSYSDHIFERDLHSLVTRA 300
DB 241 SPGSPMEITDAEKQWILNGLIEGNPEFGKNGFOLVRSYSDHIFERDLHSLVTRA 300
QY 301 GNGVYITSGDNGQSGITQKSGIPSEIKITILANMSILPKERKQYHNPYDGNITSPRLN 360
DB 301 GNGVYITSGDNGQSGITQKSGIPSEIKITILANMSILPKERKQYHNPYDGNITSPRLN 360
QY 301 GNGVYITSGDNGQSGITQKSGIPSEIKITILANMSILPKERKQYHNPYDGNITSPRLN 360
DB 301 GNGVYITSGDNGQSGITQKSGIPSEIKITILANMSILPKERKQYHNPYDGNITSPRLN 360
QY 361 NGETLYEMDQKQSLIFASBINOGAGLYFEAGNFTVSNSTQTMQAGIHVSSENSTYTWK 420
DB 361 NGETLYEMDQKQSLIFASBINOGAGLYFEAGNFTVSNSTQTMQAGIHVSSENSTYTWK 420
QY 361 NGETLYEMDQKQSLIFASBINOGAGLYFEAGNFTVSNSTQTMQAGIHVSSENSTYTWK 420
DB 361 NGETLYEMDQKQSLIFASBINOGAGLYFEAGNFTVSNSTQTMQAGIHVSSENSTYTWK 420
QY 421 VNGVGHRLSKIGKGLHVAOKGENKGSISVGGKAYILEGADQDQGNKQAFSEIGLYSR 480
DB 421 VNGVGHRLSKIGKGLHVAOKGENKGSISVGGKAYILEGADQDQGNKQAFSEIGLYSR 480
QY 421 VNGVGHRLSKIGKGLHVAOKGENKGSISVGGKAYILEGADQDQGNKQAFSEIGLYSR 480
DB 421 VNGVGHRLSKIGKGLHVAOKGENKGSISVGGKAYILEGADQDQGNKQAFSEIGLYSR 480
QY 481 GTVQIANDKQFDTKKYPFGFRGLDLNGHSLFTFKRIQNTDEGAMVNTTQAAVNTT 540
DB 481 GTVQIANDKQFDTKKYPFGFRGLDLNGHSLFTFKRIQNTDEGAMVNTTQAAVNTT 540
QY 479 GTVQIANDKQFDTKKYPFGFRGLDLNGHSLFTFKRIQNTDEGAMVNTTQAAVNTT 540
DB 479 GTVQIANDKQFDTKKYPFGFRGLDLNGHSLFTFKRIQNTDEGAMVNTTQAAVNTT 540
QY 541 GNESTVLPNGNNINKIDYKELIYVNGFGRTDKANKANRLALYKPTEDRLLISGGTN 600
DB 541 GNESTVLPNGNNINKIDYKELIYVNGFGRTDKANKANRLALYKPTEDRLLISGGTN 600
QY 539 GSDTI---NDNTGDIITNRDIAFNNGFEGKQDITNTRNLVNTPLNDKMFLLISGNT 594
DB 539 GSDTI---NDNTGDIITNRDIAFNNGFEGKQDITNTRNLVNTPLNDKMFLLISGNT 594
QY 601 LKGDITCTKGLFFSGRPTPHAYNHLKMKSEWEGIPQGBIWMDEHINRTKAEKFIK 660
DB 601 LKGDITCTKGLFFSGRPTPHAYNHLKMKSEWEGIPQGBIWMDEHINRTKAEKFIK 660
QY 595 LKGNITDQGLIVFSRPTPHAYNHLK-RLNEL-GRPKGVVTDQWIKIRTKAANPQIK 652
DB 595 LKGNITDQGLIVFSRPTPHAYNHLK-RLNEL-GRPKGVVTDQWIKIRTKAANPQIK 652
QY 661 GGSAAVSRVSSIEGNMTVSNANATGVVPMQONTICTRSDMTGLTLOKVDLITKYI 720
DB 661 GGSAAVSRVSSIEGNMTVSNANATGVVPMQONTICTRSDMTGLTLOKVDLITKYI 720
QY 653 GGSAAVSRVSSIEGNMTVSNANATGVVPMQONTICTRSDMTGLTLOKVDLITKYI 712
DB 653 GGSAAVSRVSSIEGNMTVSNANATGVVPMQONTICTRSDMTGLTLOKVDLITKYI 712

```

```

QY 721 NSIPKTOINGSINTDMATANKGLAKUNGVLTVNHSQFTISNNATQIGNILSPNSTA 780
DB 721 NSIPKTOINGSINTDMATANKGLAKUNGVLTVNHSQFTISNNATQIGNILSPNSTA 780
QY 713 DSIPKTOINGSINTDMATANKGLAKUNGVLTVNHSQFTISNNATQIGNILSPNSTA 772
DB 713 DSIPKTOINGSINTDMATANKGLAKUNGVLTVNHSQFTISNNATQIGNILSPNSTA 772
QY 781 TVDNANLNGVHLLDSKQSLKNSHFSHQSGKGTTLVNAWTWPSPTTQNTLTLNN 840
DB 781 TVDNANLNGVHLLDSKQSLKNSHFSHQSGKGTTLVNAWTWPSPTTQNTLTLNN 840
QY 773 TVDNANLNGVHLLDSKQSLKNSHFSHQSGKGTTLVNAWTWPSPTTQNTLTLNN 832
DB 773 TVDNANLNGVHLLDSKQSLKNSHFSHQSGKGTTLVNAWTWPSPTTQNTLTLNN 832
QY 841 STTILNSAVSASNTPR-RRLSETEPTPSAEHRENTLTVNGKLSGGQTFQPTSSLPFY 899
DB 841 STTILNSAVSASNTPR-RRLSETEPTPSAEHRENTLTVNGKLSGGQTFQPTSSLPFY 899
QY 833 STTILNSAVSASNTPR-RRLSETEPTPSAEHRENTLTVNGKLSGGQTFQPTSSLPFY 892
DB 833 STTILNSAVSASNTPR-RRLSETEPTPSAEHRENTLTVNGKLSGGQTFQPTSSLPFY 892
QY 900 KSDKILSNDABEGDYLISVRNTGKEPTELTQTLVSKNOQLSKKLTLENDHDAEA 959
DB 900 KSDKILSNDABEGDYLISVRNTGKEPTELTQTLVSKNOQLSKKLTLENDHDAEA 959
QY 893 KSDKILSNDABEGDYLISVRNTGKEPTELTQTLVSKNOQLSKKLTLENDHDAEA 952
DB 893 KSDKILSNDABEGDYLISVRNTGKEPTELTQTLVSKNOQLSKKLTLENDHDAEA 952
QY 960 LRYKLVKNDGFRFLNPDIKEOEIENDLYPRAQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1019
DB 960 LRYKLVKNDGFRFLNPDIKEOEIENDLYPRAQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1019
QY 953 WRTELKNGGFRFLNPDIKEOEIENDLYPRAQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1010
DB 953 WRTELKNGGFRFLNPDIKEOEIENDLYPRAQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1010
QY 1020 ARAAPPDTLPDQSLNLEAKQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1064
DB 1020 ARAAPPDTLPDQSLNLEAKQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1064
QY 1011 -RAESDTPPQSQNLNLEAKQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1069
DB 1011 -RAESDTPPQSQNLNLEAKQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1069
QY 1065 FSDPLDQSLNLEAKQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1108
DB 1065 FSDPLDQSLNLEAKQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1108
QY 1070 LATALVEKTAQIDPANAALQNLNLEAKQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1126
DB 1070 LATALVEKTAQIDPANAALQNLNLEAKQAEFTLEAKQVEPFAKTQTEGPEVRSRA 1126
QY 1109 LISRYNSALSELATVNSMLSVODELRLPVDQAQSAVNTIADKKRYDSDAFAYQO 1168
DB 1109 LISRYNSALSELATVNSMLSVODELRLPVDQAQSAVNTIADKKRYDSDAFAYQO 1168
QY 1127 LISRYNSALSELATVNSMLSVODELRLPVDQAQSAVNTIADKKRYDSDAFAYQO 1185
DB 1127 LISRYNSALSELATVNSMLSVODELRLPVDQAQSAVNTIADKKRYDSDAFAYQO 1185
QY 1169 OKTNRLQIGVQKALANGRIQAVFSHRSNDTEDEQVKNATLTMNSGPAQYQNGDLOFQV 1228
DB 1169 OKTNRLQIGVQKALANGRIQAVFSHRSNDTEDEQVKNATLTMNSGPAQYQNGDLOFQV 1228
QY 1186 OKTNRLQIGVQKALANGRIQAVFSHRSNDTEDEQVKNATLTMNSGPAQYQNGDLOFQV 1245
DB 1186 OKTNRLQIGVQKALANGRIQAVFSHRSNDTEDEQVKNATLTMNSGPAQYQNGDLOFQV 1245
QY 1229 NVGTGTSASKAEKQSKIRKALVGNASYSFHLQGLDIPYGVNRYTERKRYOSE 1288
DB 1229 NVGTGTSASKAEKQSKIRKALVGNASYSFHLQGLDIPYGVNRYTERKRYOSE 1288
QY 1246 NVGTGTSASKAEKQSKIRKALVGNASYSFHLQGLDIPYGVNRYTERKRYOSE 1305
DB 1246 NVGTGTSASKAEKQSKIRKALVGNASYSFHLQGLDIPYGVNRYTERKRYOSE 1305
QY 1289 EYVKTPTSLAEFRNAGIRVDYFTPTDNISVKEPFPVNVDSNANQVTLNLTVLQOP 1348
DB 1289 EYVKTPTSLAEFRNAGIRVDYFTPTDNISVKEPFPVNVDSNANQVTLNLTVLQOP 1348
QY 1306 EYVKTPTSLAEFRNAGIRVDYFTPTDNISVKEPFPVNVDSNANQVTLNLTVLQOP 1365
DB 1306 EYVKTPTSLAEFRNAGIRVDYFTPTDNISVKEPFPVNVDSNANQVTLNLTVLQOP 1365
QY 1349 EYVKTPTSLAEFRNAGIRVDYFTPTDNISVKEPFPVNVDSNANQVTLNLTVLQOP 1394
DB 1349 EYVKTPTSLAEFRNAGIRVDYFTPTDNISVKEPFPVNVDSNANQVTLNLTVLQOP 1394
QY 1366 EYVKTPTSLAEFRNAGIRVDYFTPTDNISVKEPFPVNVDSNANQVTLNLTVLQOP 1411
DB 1366 EYVKTPTSLAEFRNAGIRVDYFTPTDNISVKEPFPVNVDSNANQVTLNLTVLQOP 1411

```

```

RESULT 9
US-10-687-046-17
; Sequence 17, Application US/10687046
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/687,046
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US/10/080,505
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-687-046-17

```

Query Match 80.1%; Score 5827.5; DB 31; Length 1411;
 Best Local Similarity 79.5%; Pred. No. 0; Mismatches 139; Indels 47; Gaps 12;
 Matches 1133; Conservative 107;

QY 1 MKKTVRLNPLFLACISLGIYSQAMAGHTYFGIDYQYRDPANRKGKFTVGAONIKYVNRKQ 60
 DB 1 MKKTVRLNPLFLACISLGIYSQAMAGHTYFGIDYQYRDPANRKGKFTVGAONIKYVNRKQ 60

QY 61 GQVGTSMTKAPMIDFSVSRNGVAALYENQYIVSAHNNGTVDPAAGANNDDHRT 120
 DB 61 GQVGTSMTKAPMIDFSVSRNGVAALYENQYIVSAHNNGTVDPAAGANNDDHRT 120

QY 121 YKIVRNNYKKONLHPYEDDYHNPRLHKFTLEAPIDMTSNMNGSTYSDRKKPERVRIG 180
 DB 121 YKIVRNNYKKONLHPYEDDYHNPRLHKFTLEAPIDMTSNMNGSTYSDRKKPERVRIG 180

QY 181 SGRQWRNDQKQDVAGAVHYLLTAGNTHNORGANGSYLGGDVRKAGEYGLPLAGSK 240
 DB 181 SGRQWRNDQKQDVAGAVHYLLTAGNTHNORGANGSYLGGDVRKAGEYGLPLAGSK 240

QY 241 GDSGSPMFTYDAEKOKMLNGILLRBNPBEKENGFOJVRKSYFDEIFERDLTSLYTRA 300
 DB 241 GDSGSPMFTYDAEKOKMLNGILLRBNPBEKENGFOJVRKSYFDEIFERDLTSLYTRA 300

QY 301 GNGVITISGNDGSGSITOKSGIPSEIKITLANNSLPKKOKYHNPYDGENYSEPLN 360
 DB 301 GNGVITISGNDGSGSITOKSGIPSEIKITLANNSLPKKOKYHNPYDGENYSEPLN 360

QY 361 NGEITLYMDOKQSLIFASDINOAGGLYFESGFTVSPNSNQTWQAGIHVENSJYTWK 420
 DB 361 NGEITLYMDOKQSLIFASDINOAGGLYFESGFTVSPNSNQTWQAGIHVENSJYTWK 420

QY 421 VNGVEHRLSKIGKTLHVOAKGKNGSISVBDGKYLLEQADDOQAASEIGLVSR 480
 DB 421 VNGVEHRLSKIGKTLHVOAKGKNGSISVBDGKYLLEQADDOQAASEIGLVSR 480

QY 481 GTVOLNDDKQEDTKEFESRGRDLNNGHSLTKRLONTDEGAMVNHNTQOANYTT 540
 DB 481 GTVOLNDDKQEDTKEFESRGRDLNNGHSLTKRLONTDEGAMVNHNTQOANYTT 540

QY 541 GNEISIVLPNGNNINKLTYRKEIAYNGMGEJDKNKNRNLALTYPTEDRTLLSGGTN 600
 DB 541 GNEISIVLPNGNNINKLTYRKEIAYNGMGEJDKNKNRNLALTYPTEDRTLLSGGTN 600

QY 601 LKEDITQTKGLFESGRPTPHAYNHLNKRMSMGEISQGIWDDHMLNRTFKAKNFQIK 660
 DB 601 LKEDITQTKGLFESGRPTPHAYNHLNKRMSMGEISQGIWDDHMLNRTFKAKNFQIK 660

QY 661 GGSAYVSRVSSIEGNTMTVSNNANATFGVVPQONTCTRSWGTGLTCCOKVDELTKYI 720
 DB 661 GGSAYVSRVSSIEGNTMTVSNNANATFGVVPQONTCTRSWGTGLTCCOKVDELTKYI 720

QY 721 NSIPKQINGSLNLTNANATVYKGLAKLKNVTLTNHSGFTLSNNATQIGNRLSDNSTA 780
 DB 721 NSIPKQINGSLNLTNANATVYKGLAKLKNVTLTNHSGFTLSNNATQIGNRLSDNSTA 780

QY 781 TVDNANLNGNVAHLTSAQFSLKNSHSHQIQDCKTIVYLENATWMSOFTLONLTAN 840
 DB 781 TVDNANLNGNVAHLTSAQFSLKNSHSHQIQDCKTIVYLENATWMSOFTLONLTAN 840

QY 841 STTINSAYSASANNTPR-PRSLKETPTPTSHERFNTLVNKGKSGOCTQFQPSLFG 899
 DB 841 STTINSAYSASANNTPR-PRSLKETPTPTSHERFNTLVNKGKSGOCTQFQPSLFG 899

QY 899 KSDTKLNSDADGDIYLSVNTGKEBELTQJLVESKDNQPLSDKFLTLNDHVDAGA 959
 DB 899 KSDTKLNSDADGDIYLSVNTGKEBELTQJLVESKDNQPLSDKFLTLNDHVDAGA 959

QY 959 LRYKLYVNGEERLHNPILKQELHNDLVABAEKTLIAQVAFPAKYOTGEPPYRERRA 1019
 DB 959 LRYKLYVNGEERLHNPILKQELHNDLVABAEKTLIAQVAFPAKYOTGEPPYRERRA 1019

QY 1019 WRZEIVKXGGEERLHNPILKQELHNDLVABAEKTLIAQVAFPAKYOTGEPPYRERRA 1019
 DB 1019 WRZEIVKXGGEERLHNPILKQELHNDLVABAEKTLIAQVAFPAKYOTGEPPYRERRA 1019

QY 1020 ARAAFPTLIPDQSLINLLEAKQAELEAFQSKA-----KTKKVSER- - - - -AV 1064
 DB 1020 ARAAFPTLIPDQSLINLLEAKQAELEAFQSKA-----KTKKVSER- - - - -AV 1064

DB 1011 -RAAFSTDPDQSLINLLEAKQAELEAFQSKA-----KTKKVSER- - - - -AV 1064
 QY 1065 FSDPLDQSLINLLEAKQAELEAFQSKA-----KTKKVSER- - - - -AV 1064
 DB 1070 LATALEKETAQIDFANAKLAQMLTQOLEKALAV---AEQAEKERRAEOAKRQKQD 1126
 QY 1109 LLSRYSNSALSELATVNSMCLSYDELDRLFYDQASAVNTNIOKDRYDPAFAYQ 1166
 DB 1127 LLSRYSNSALSELATVNSMCLSYDELDRLFYDQASAVNTNIOKDRYDPAFAYQ 1166

QY 1169 OKTNLROIGVQKALANGRIGAVFSHRSKSDTPEQVKNATLTMGSGPAQYQMGDLQFQV 1228
 DB 1186 QCTNLRQIGVQKALANGRIGAVFSHRSKSDTPEQVKNATLTMGSGPAQYQMGDLQFQV 1245

QY 1229 NVGTGISASMAEGRKTHRRALNYGNVNSYQFRLQGLIOPFYGNRYFIERENYQSE 1288
 DB 1246 NVGTGISASMAEGRKTHRRALNYGNVNSYQFRLQGLIOPFYGNRYFIERENYQSE 1305

QY 1289 ERYKVTBSLAPNRYNAGRVDTFTPTDNTSVKPYFNNYVDVSNNAVOTTVNLTVLQOP 1348
 DB 1306 ERYKVTBSLAPNRYNAGRVDTFTPTDNTSVKPYFNNYVDVSNNAVOTTVNLTVLQOP 1365

QY 1349 FERYQKEVGIKAEILHFPQISAFISKSQSGQLGKQONVYKLYRW 1394
 DB 1366 FERYQKEVGIKAEILHFPQISAFISKSQSGQLGKQONVYKLYRW 1411

RESULT 10
 PCT-US03-05226-13
 ; Sequence 13, Application PC/TUS0305226
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RT/DC/DBR
 ; CURRENT APPLICATION NUMBER: PCT/US03/05226
 ; CURRENT FILING DATE: 2003-02-18
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 1436
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; PCT-US03-05226-13

Query Match 79.8%; Score 5807; DB 1; Length 1436;
 Best Local Similarity 79.3%; Pred. No. 0;
 Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;

QY 1 MKKTVRLNPLFLACISLGIYSQAMAGHTYFGIDYQYRDPANRKGKFTVGAONIKYVNRKQ 60
 DB 1 MKKTVRLNPLFLACISLGIYSQAMAGHTYFGIDYQYRDPANRKGKFTVGAONIKYVNRKQ 60

QY 61 GQVGTSMTKAPMIDFSVSRNGVAALYENQYIVSAHNNGTVDPAAGANNDDHRT 120
 DB 61 GQVGTSMTKAPMIDFSVSRNGVAALYENQYIVSAHNNGTVDPAAGANNDDHRT 120

QY 121 YKIVRNNYKKONLHPYEDDYHNPRLHKFTLEAPIDMTSNMNGSTYSDRKKPERVRIG 180
 DB 121 YKIVRNNYKKONLHPYEDDYHNPRLHKFTLEAPIDMTSNMNGSTYSDRKKPERVRIG 180

QY 181 SGRQWRNDQKQDVAGAVHYLLTAGNTHNORGANGSYLGGDVRKAGEYGLPLAGSK 240
 DB 181 SGRQWRNDQKQDVAGAVHYLLTAGNTHNORGANGSYLGGDVRKAGEYGLPLAGSK 240

QY 241 GDSGSPMFTYDAEKOKMLNGILLRBNPBEKENGFOJVRKSYFDEIFERDLTSLYTRA 300
 DB 241 GDSGSPMFTYDAEKOKMLNGILLRBNPBEKENGFOJVRKSYFDEIFERDLTSLYTRA 300

QY 295 SLTYRAGNGVTTISGNDGQSSITQ---KSGIPSEIKITLANNSLPEKXKHNPY-- 349
 DB 301 TELPRNSNGHSTSNNGGCTVYQNEKYSWPOFYKRYTQULNEALKEKXK--EPYVA 358
 QY 350 DGPVYSPRLNNGETLYFMDQKQSLIFASDINOGAGGLYFEGNPFVSPNSNOTWOGAGI 409
 DB 359 GGVNAYKPRLNNGKNIYFGDRGTITLITENINOGAGGLYFEGNPFVSENNATWOGAGV 418
 QY 410 HVSNSVTVMKNGVEHDLRSKIGKGLTFVQAKGENKSI SVGDQVYI..EQOADDQGNQ 469
 DB 419 HVBGDSVTVMKNGVEHDLRSKIGKGLTFVQAKGENKSI SVGDQVYI..EQOADDQGNQ 478
 QY 470 AFSEIGLVSGGTQVQANDKQFDTDKFYFGSGGGLDNGHSLTFRIGTQNTDEGAMTVNH 529
 DB 479 AFKEGVIVSGRATVQJNSADQVDPNNIYFGFGRGLDNGHSLTFRIGTQNTDEGAMTVNH 538
 QY 530 NTQOANVTITGNES-VLP-NGNNINKLDRYREILAYNGWFEGEDKXENGRNLNLYEPT 588
 DB 539 NTQOANVTITGNES-VLP-NGNNINKLDRYREILAYNGWFEGEDKXENGRNLNLYEPT 598
 QY 589 EDRLLISGGTNLKGDIYQTKGLFSGGRPTPHAYNHLAKRWSMEGIPOGRIVWDHWT 648
 DB 599 EDRLLISGGTNLKGDIYQTKGLFSGGRPTPHAYNHLAKRWSMEGIPOGRIVWDHWT 656
 QY 649 NRTKANEQIKGSNAVSRVSSIEGNMTVSNNAKAFGVYVPOQNTICRSDMTGLTT 708
 DB 657 TRTKANEQIKGSNAVSRVSSIEGNMTVSNNAKAFGVYVPOQNTICRSDMTGLTT 716
 QY 709 CQVYDLTDTKINS:FKTOINGSINTLDAATNWKGLAKINGVTLTINHSQFTSNNATQ 768
 DB 717 CKYVTLDTKIVINS:FTQOINGSINTLDAATNWKGLAKINGVTLTINHSQFTSNNATQ 776
 QY 769 IGNIRLSDNSTATVNDANLNGVHLDTSAQSLKXSHESHQISQDGGTITVLENATVMP 828
 DB 777 TGNITLSDNSTATVNDANLNGVHLDTSAQSLKXSHESHQISQDGGTITVLENATVMP 836
 QY 829 SDPTLQNTLINS:TLNSAYSASNNTPR-RSEFTTPTTSAHERPHTLVNGLXISGQ 887
 DB 837 SDATLQNTLINS:TLNSAYSASNNTPR-RSEFTTPTTSAHERPHTLVNGLXISGQ 896
 QY 888 STPOFTSS:FGYKSDKLKLSNDAEGVYILSVNTEKEPETLEQTLVESKONPISDKL 947
 DB 897 STPOFTSS:FGYKSDKLKLSNDAEGVYILSVNTEKEPETLEQTLVESKONPISDKL 956
 QY 946 FTLENDHVDAGALRYKLVNDGEFRLANPIKEQSLHNDLVPAQERTEAKOYBPFAKT 1007
 DB 957 FTLENDHVDAGALRYKLVNDGEFRLANPIKEQSLHNDLVPAQERTEAKOYBPFAKT 1016
 QY 1008 QTGPRKYSRRAARAAPBPTLPDQSLINLMEAKOELFAETQSKAKTKQKSRRAVPSD 1067
 DB 1017 QTSNARVRSK--RAVFSDTLPDQSLQVQAEVPEPTAEKOKNKA--EKVASKRAVPSD 1071
 QY 1068 FLDOS-----LFALEALEVID 1085
 DB 1072 TLTPQSQDLVQAEVPEPTAEKOKNKAAYKRSKRAAREPSTPLDLSKLYEVLKLVIN 1131
 QY 1086 APQOSEKORLAQEAERKQKXLIISYNSALSTLSATVNSM:SVODELRLPYDQAS 1145
 DB 1133 AQOQVKE--POQEKORQKXLIISYNSALSTLSATVNSM:SVODELRLPYDQAS 1188
 QY 1146 AVMTNIAQDRAVDSAPRAYQOQKNTLQICQKALMNGIAGVSHRSRNTFDDQV 1205
 DB 1189 AVMTNIAQDRAVDSAPRAYQOQKNTLQICQKALMNGIAGVSHRSRNTFDDQV 1247
 QY 1206 NHATLTMSSGFAQYQMGDLOFQVNVGTGISASMAEBSRKIRKAINGVNAYSORLQ 1265
 DB 1248 NHATLTMSSGFAQYQMGDLOFQVNVGTGISASMAEBSRKIRKAINGVNAYSORLQ 1307
 QY 1266 QLGIOFYFGNARFIERENYQSEEVYKTPSLAEFRNAGIRUDYPTTNDISIKYEPF 1325
 DB 1308 QLGIOFYFGNARFIERENYQSEEVYKTPSLAEFRNAGIRUDYPTTNDISIKYEPF 1367
 QY 1326 VNYVDVSNANVQTTNVLVLOQFGRYWKVEGLAKAELIHFQISATISKSQSQJAKQON 1385

DB 1368 VNYVDVSNANVQTTNVLVLOQFGRYWKVEGLAKAELIHFQISATISKSQSQJAKQON 1427
 QY 1386 VGVXKGYRW 1394
 DB 1428 VGVXKGYRW 1436
 RESULT 11
 US-10-687-046-13
 ; Sequence 13, Application US/10687046
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OR INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RET/DCF/DER
 ; CURRENT APPLICATION NUMBER: US/10/687, 046
 ; PRIOR FILING DATE: 2003-10-15
 ; PRIOR APPLICATION NUMBER: US/10/080, 505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296, 791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839, 996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 1436
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-10-687-046-13
 Query Match 79.8%; Score 5807; DB 31; Length 1436;
 Best Local Similarity 79.3%; Pred. No. 0;
 Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;
 QY 1 MKKTVFRLNPLTACISIGIVSOAMAGTYRGIDYQYRDPFAENKSEFTVGAQNIKYVNO 60
 DB 1 MKKTVFRLNPLTACISIGIVSOAMAGTYRGIDYQYRDPFAENKSEFTVGAQNIKYVNO 60
 QY 61 GQIVGTSMTKAPMIDSVSVSRNGVAALVENQYIVSAHVNGYTTDVPQAGSNPDDHRT 123
 DB 61 GQIVGTSMTKAPMIDSVSVSRNGVAALVENQYIVSAHVNGYTTDVPQAGSNPDDHRT 120
 QY 121 YKIVRNRYKKNLHPEDDYRNPRHKEVYTEAPIDMTSNNGSYSDSKYBESVRIG 180
 DB 121 YQIVRNRYKKNLHPEDDYRNPRHKEVYTEAPIDMTSNNGSYSDSKYBESVRIG 180
 QY 181 SGROFPRNDQD-----KQDVAGAYHYLLTANTHORGAGNGSYLGGDYKAKAGEYGP 235
 DB 181 TSGQWKTDEQKQSGKSSMLADATVLMILGNTHSGAGNGTVNLSGDITKNNYGP 240
 QY 236 IAGSKGDSGSPFTYDAEYQKWLINGILRSGNPPBKGKNGFQYVRSYF-DEIFERDLHT 294
 DB 241 TGVSFEDSGSPFTYDAEYQKWLINGVLTQNGNPFSGAGNGFQILIRKWFYDVNFVEDLEI 300
 QY 295 SLTYRAGNGVTTISGNDGQSSITQ---KSGIPSEIKITLANNSLPEKXKHNPY-- 349
 DB 301 TELPRNSNGHSTSNNGGCTVYQNEKYSWPOFYKRYTQULNEALKEKXK--EPYVA 358
 QY 350 DGPVYSPRLNNGETLYFMDQKQSLIFASDINOGAGGLYFEGNPFVSPNSNOTWOGAGI 409
 DB 359 GGVNAYKPRLNNGKNIYFGDRGTITLITENINOGAGGLYFEGNPFVSENNATWOGAGV 418
 QY 410 HVSNSVTVMKNGVEHDLRSKIGKGLTFVQAKGENKSI SVGDQVYI..EQOADDQGNQ 469
 DB 419 HVBGDSVTVMKNGVEHDLRSKIGKGLTFVQAKGENKSI SVGDQVYI..EQOADDQGNQ 478
 QY 470 AFSEIGLVSGGTQVQANDKQFDTDKFYFGSGGGLDNGHSLTFRIGTQNTDEGAMTVNH 529
 DB 479 AFKEGVIVSGRATVQJNSADQVDPNNIYFGFGRGLDNGHSLTFRIGTQNTDEGAMTVNH 538
 QY 530 NTQOANVTITGNES-VLP-NGNNINKLDRYREILAYNGWFEGEDKXENGRNLNLYEPT 588
 DB 539 NTQOANVTITGNES-VLP-NGNNINKLDRYREILAYNGWFEGEDKXENGRNLNLYEPT 598

Db 539 NITQVANITITNESITAPSNKNNIKLDYSKEIAVNGFGETDENKNGRLNIYKPTT 598
 QY 565 EDETLISGNTIKGDIITQKGLFSSGPPFAHNLAKESMEGIPQGEIYWHMI 648
 Db 593 EDETLISGNTIKGDIITQKGLFSSGPPFAHNLAKESMEGIPQGEIYWHMI 656
 QY 649 NRTFKANFOIKGSAVAVSRVSSIEGNTWTSNNANATFGVNPQONTICRSPTMTGLT 708
 Db 657 TRTFKANKFOIKGSAVAVSRVSSIEGNTWTSNNANATFGVNPQONTICRSPTMTGLT 716
 QY 709 CQKVDITDKVINSIPTQINGSLNLTNATANYGLAKINGVTLNHSQFTLSNNATQ 768
 Db 717 CKTVLITDKVINSIPTQINGSLNLTNATANYGLAKINGVTLNHSQFTLSNNATQ 776
 QY 769 IGNIKLSDNSTATVDNANLNGVNLTDASQSLKNSHFSHQIGQKGTVTTLNATWMP 828
 Db 777 TGNILSNANATVDNANLNGVNLTDASQSLKNSHFSHQIGQKGTVTTLNATWMP 836
 QY 829 SDPTLQNTLANSITLINSAYSSASNNPR-RSE-ETETPTSAEHRNTLTVNGLSQ 887
 Db 837 SDPTLQNTLANSITLINSAYSSASNNAPRRKSLTETETPTSAEHRNTLTVNGLSQ 896
 QY 888 GTFPTSLIFGYSKDKLKSLNDAEGDYLLSVRNTGKPELTLEQLTVESKDNOPLSDKLK 947
 Db 897 GTFPTSLIFGYSKDKLKSLNDAEGDYLLSVRNTGKPELTLEQLTVESKDNOPLSDKLK 956
 QY 948 FTLNHDVDAALRYKLVKNDPEPLNHPITEGELNDIVRAEQAEITLAKVETPACT 1007
 Db 957 FTLNHDVDAALRYKLVKNDPEPLNHPITEGELNDIVRAEQAEITLAKVETPACT 1016
 QY 1008 QTGEPRVSRRAAFAEPDILPDOSILNLEAKQAEITLAKVETPACT 1067
 Db 1017 QTSNARVSK---RAVSPDILPDOSILNLEAKQAEITLAKVETPACT 1071
 QY 1068 PLIDQS-----FALSALEVID 1085
 Db 1072 TLPDOSQDLVLAQVEYFLAEKQKNAKAVRSKRAAREPSPDILSKLAVLEVID 1131
 QY 1086 APOSEKORLAQEAERKOKKODILSRYSNALSSELATVMSMLSYODEILDFLPDQOS 1145
 Db 1132 APOQVKE---PDQSEKORQODILSRYSNALSSELATVMSMLSYODEILDFLPDQOS 1188
 QY 1146 AVWTIAQDKRKYSDAFRAVQOQKTNLRQIGOKALANGRIAGVFSHSNDTPEQYK 1205
 Db 1189 AVWTIAQDKRKYSDAFRAVQOQKTNLRQIGOKALANGRIAGVFSHSNDTPEQYK 1247
 QY 1206 NHATITMWSGPAQYQWDLQFQSVNVTGISASMAEQRKIHRAINYGNVANSYQFRLG 1265
 Db 1248 NHATITMWSGPAQYQWDLQFQSVNVTGISASMAEQRKIHRAINYGNVANSYQFRLG 1307
 QY 1266 QGIGOPYGVNRYFIRENYSSEBRVKTSLAFNRYNAGRINTYTTPTDNI.SVYPTFF 1325
 Db 1308 QGIGOPYGVNRYFIRENYSSEBRVKTSLAFNRYNAGRINTYTTPTDNI.SVYPTFF 1367
 QY 1326 VNYVDVSNANQTTVALTVLQOPFGRYWKQEVGLKELIHFQISAFISKSQSGQKQON 1385
 Db 1368 VNYVDVSNANQTTVALTVLQOPFGRYWKQEVGLKELIHFQISAFISKSQSGQKQON 1427
 QY 1386 VGVKLYRW 1394
 Db 1428 VGVKLYRW 1436

RESULT 12
 PCT-US03-05226-15
 ; Sequence 15, Application PC/TUS0305226
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; CURRENT APPLICATION NUMBER: PCT/US03/05226
 ; CURRENT FILING DATE: 2003-02-18
 ; PRIOR APPLICATION NUMBER: US 08/256,791

; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 1391
 ; TYPE: EXT
 ; ORGANISM: Haemophilus influenzae
 ; PCT-US03-05226-15

Query Match 72.8%; Score 5303.5; DB 1; Length 1391;
 Best Local Similarity 74.8%; Pred. No. 0;
 Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

QY 1 MKKTVPLNLEFLACISLIGVSOANAAGTYGIDYQYDFAPENKGRPTVGAQNIKYVKKQ 60
 Db 1 MKKTVPLNLEFLACISLIGVSOANAAGTYGIDYQYDFAPENKGRPTVGAQNIKYVKKQ 60
 QY 61 GQLVGTSMTAEMIDFVSYSRNGVAAVENVQYIVSAHVAGYTDVDPGAEQNNPDQHRFT 120
 Db 61 GEMIGTMKGVPMWPLDSNWRGYSYLISHQHLISVAHVAGYDVDPGAEQNNPDQHRFX 120
 QY 121 YKIKRNNYKQNLHPEDDYHNRRLKRYTAAAPIDMTSMNGSTYSRRTKYPERRVIG 180
 Db 121 YKIKRNNYKQNLHPEDDYHNRRLKRYTAAAPIDMTSMNGSTYSRRTKYPERRVIG 178
 QY 121 YKIKRNNYKQNLHPEDDYHNRRLKRYTAAAPIDMTSMNGSTYSRRTKYPERRVIG 178
 Db 121 YKIKRNNYKQNLHPEDDYHNRRLKRYTAAAPIDMTSMNGSTYSRRTKYPERRVIG 178
 QY 181 SGROFMRNDQDK--GDQVAGAYHTLTAGTNHQRQAGNGSYLGGDVRKAGEYGLPIAG 238
 Db 179 SGHGMWMDQNNKTIDILAVYG--SWLIGNTPEPDGAGNGTLELNGVQNPNNYGLPIAG 237
 QY 239 SKQSGSPMEFYDAKQKYLINGLREBQNPFEKNGFQVLRKSYDEILPEEDLHTSLYT 298
 Db 238 SFGDGSQPMFYDKKQKYLINGLREBQNPFEKNGFQVLRKSYDEILPEEDLHTSLYT 297
 QY 299 RAGNGVYTTISGNDG-----QGSITQSGIPEITITLANSI--PLKEDCKVHNRDGP 352
 Db 298 TNAETFRNIGSDHNRVATITSTLPKKAQPERIVGLVNSQLHARDNRNGDSESYKGP 357
 QY 353 NIVSPRLNNGETLYEMDQKQSLIFASDINGAGGLYEPGNFTVSPNSQ--TMQAGIHY 411
 Db 358 NPMSPALHGHKSLYRGDQGTGLTIENNINQAGGLYEPGNFTVSPNSQ--TMQAGIHY 417
 QY 412 SENSTYTWKNGVEHDLRSLKIGKTLHVQAKENKGSISYQKQVTLLEQADQDQKQAF 471
 Db 418 GBSSTYTWKNGVEHDLRSLKIGKTLHVQAKENKGSISYQKQVTLLEQADQDQKQAF 477
 QY 472 SEIGLVSGRTVQANDQKQFPTQKPYFGRGRDLNGHSLTEKRIQNTDQBGAMTVNHT 531
 Db 478 KEVGIYSGRATVQANSADQVDPNNIYFSPRGRLDLNGHSLTEKRIQNTDQBGAMTVNHT 537
 QY 532 TOAANVTITNESIVLPNGNNINKLDYRKETAVNGMFGETDKKNGRNLTIYPTTEDR 591
 Db 538 SOTANITITGNATI---NSDSKQILNKDILAPGCMFGEDQKAKTNGRNLNVPVAEN 593
 QY 592 TLLISGNTUKDITQKGLFSSGPPFAHNLAKESMEGIPQGEIYWHDMKNT 651
 Db 594 HLLISGNTUKDITQKGLFSSGPPFAHNLAKESMEGIPQGEIYWHDMKNT 653
 QY 652 FKAEKFOIKGSAVAVSRVSSIEGNTWTSNNANATFGVNPQONTICRSPTMTGLTCTQK 711
 Db 654 FKAEKFOIKGSAVAVSRVSSIEGNTWTSNNANATFGVNPQONTICRSPTMTGLTCTK 713
 QY 712 VDLITDKVINSIPTQINGSLNLTNATANYGLAKINGVTLNHSQFTLSNNATQIGN 771
 Db 714 VDLITDKVINSIPTQINGSLNLTNATANYGLAKINGVTLNHSQFTLSNNATQIGN 773
 QY 772 IRLSDNSTATVDNANLNGVNLTDASQSLKNSHFSHQIGQKGTVTTLNATWMP 831
 Db 774 IRLSDNSTATVDNANLNGVNLTDASQSLKNSHFSHQIGQKGTVTTLNATWMP 833
 QY 832 TLQNTLANSITLINSAYSSASNNPR--RRSLTETPTSAEHRNTLTVNGLSQSGGT 889

Db 834 TLONLTLNNSVTLSAAYSALSINAPRRRSLEFETTPPSASHPNTLTUNGKLSGGCT 893
 Qy 890 FQFTSSLFGYKSDKLTLSNDAEGDYLLSVRNTGKEPBTLPQTLTVESKDNQPLSDKLTFT 949
 Db 894 FQFTSSLFGYKSDKLTLSNDAEGDYLLSVRNTGKEPBTLPQTLTVESKDNQPLSDKLTFT 953
 Qy 950 LENDVDAGALRYKLVKNDGEPRLNPIKEGSLHNDLVRAEQERTLEAKQVEQTACTOT 1009
 Db 954 LENDVDAGALRYKLVKNDGEPRLNPIKEGSLHNDLVRAEQERTLEAKQVEQTACTOT 1013
 Qy 1010 GEPKRSRRAARAAPDPLPDOSLNLALBAKAEILTAETOKSAKTKKYSKRAV--FSD 1067
 Db 1014 SKARVR--RAVSDPLPAOSLNLALBAKAEILTAETOKSAKTKKYSKRAARFSD 1067
 Qy 1068 PLDDQSLPALBAALVYIDAPQSEKRLAQBAAEK--QKQKDIISYNSALSLSATVN 1126
 Db 1068 TLDPQ--LLOALAEVYIDAQQVKEPQTOBEERKQKCKE--LSRYSNALSLSATVN 1124
 Qy 1127 SMLSVDDELDRFLVDOAASAVMTNIAODKRRYDSDAPRAYQOQKNTLRQIGYOKALNGR 1186
 Db 1125 SMLSVDDELDRFLVDOAASAVMTNIAODKRRYDSDAPRAY--QOKTLRQIGYOKALDNGR 1183
 Qy 1187 IGAVFSHSRSDNTFDEQVKNAITLTMMSGFAQYQWGLQFGVAVGTGISASMAEBSRK 1246
 Db 1184 IGAVFSHSRSDNTFDEQVKNAITLTMMSGFAQYQWGLQFGVAVGTGISASMAEBSRK 1243
 Qy 1247 IHRALVYGNASVYFRLGQLGLOPFVGNRYTIRRENYSECVKVTLSAENRYAGI 1306
 Db 1244 IHRALVYGNASVYFRLGQLGLOPFVGNRYTIRRENYSECVKVTLSAENRYAGI 1303
 Qy 1307 RVDYTFPTDNLSVYFVYVYVDSNANVOTTVNLTVLOQPGRYMOKEVGLAKETLHE 1366
 Db 1304 RVDYTFPTDNLSVYFVYVYVDSNANVOTTVNLTVLOQPGRYMOKEVGLAKETLHE 1363
 Qy 1367 QLSAFISKSQSQSLGKQONVVKLGFRM 1394
 Db 1364 QLSAFISKSQSQSLGKQONVVKLGFRM 1391
 RESULT 13
 US-10-687-046-15
 : Sequence 15, Application US/10687046
 : GENERAL INFORMATION:
 : APPLICANT: St. Gene, Joseph W.
 : TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 : FILE REFERENCE: A-59941-1/RFT/DOF/DHR
 : CURRENT APPLICATION NUMBER: US/10/687,046
 : PRIOR FILING DATE: 2003-10-15
 : PRIOR APPLICATION NUMBER: US/10/080,505
 : PRIOR FILING DATE: 2002-02-22
 : PRIOR APPLICATION NUMBER: US 08/296,791
 : PRIOR FILING DATE: 1994-10-25
 : PRIOR APPLICATION NUMBER: US 09/839,996
 : PRIOR FILING DATE: 2001-04-20
 : NUMBER OF SEQ ID NOS: 58
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 15
 : LENGTH: 1391
 : TYPE: PRF
 : ORGANISM: Haemophilus influenzae
 US-10-687-046-15
 Query Match 72.9% Score 5303.5; DB 31; Length 1391;
 Best Local Similarity 74.8% Pred. No. 0;
 Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;
 Qy 1 MKKTVFLNLTACTLSGIVSQAGHTYFGIDYQYRDPAENKGFYGAQNIKYNKQ 60
 Db 1 MKKTVFLNLTACTLSGIVSQAGHTYFGIDYQYRDPAENKGFYGAQNIKYNKQ 60
 Qy 61 GQVGTSMKAPMIDFVSVSPNGVAAIVENQYTVSAHNGYTDVDFGAGNNDQGRFT 120
 Db 61 GEMIGTMMKGVPMIDFVSVSPNGVAAIVENQYTVSAHNGYTDVDFGAGNNDQGRFT 120

Qy 221 YKIVARNYKKNLHPEEDYHNPELHKFYTEAPIDMTSNNGSYSDRTKPERVRIG 180
 Db 221 YKIVARNYKKNLHPEEDYHNPELHKFYTEAPIDMTSNNGSYSDRTKPERVRIG 178
 Qy 261 SGROFMRDQK--GQVAGAVHYITAGNTHKQKAGKSYSLGSDYRKAAGEPLPIAG 238
 Db 279 SGROFMRDQK--GQVAGAVHYITAGNTHKQKAGKSYSLGSDYRKAAGEPLPIAG 237
 Qy 239 SKGDSGPMFYDAEQKWLINGILREGNPECKEKGFLVYKSYFDEIPEERDLHTSYT 298
 Db 238 SKGDSGPMFYDAEQKWLINGILREGNPECKEKGFLVYKSYFDEIPEERDLHTSYT 297
 Qy 299 RAGNGVYITSGDNG----QGSITQSGIPSEIKITLANSI--FLKXKQKVNPRYDGP 352
 Db 298 RAGNGVYITSGDNG----QGSITQSGIPSEIKITLANSI--FLKXKQKVNPRYDGP 357
 Qy 353 NITYSPRLNGBEILYFMDQKSLIPADINQAGGIFPFGNFTYSPNSQ--TMOGAGHY 411
 Db 358 NITYSPRLNGBEILYFMDQKSLIPADINQAGGIFPFGNFTYSPNSQ--TMOGAGHY 417
 Qy 412 SENSIVTWKVGVEHRLSKICKGTLHYQAKGNKSGISVGGKVLIEQOADDQGNKQAF 471
 Db 413 SENSIVTWKVGVEHRLSKICKGTLHYQAKGNKSGISVGGKVLIEQOADDQGNKQAF 477
 Qy 472 SEIGVSGRGTVQUNDKQFDTKFTEFGGRLDNGHSLTFKRLGNTDEGAMVNT 531
 Db 478 SEIGVSGRGTVQUNDKQFDTKFTEFGGRLDNGHSLTFKRLGNTDEGAMVNT 537
 Qy 532 TQANVVTITGNESIVLPNKNINLKYKELAVNGFEGTQKKNKRNILLYKPTEDR 591
 Db 538 TQANVVTITGNESIVLPNKNINLKYKELAVNGFEGTQKKNKRNILLYKPTEDR 593
 Qy 592 TLILSGTNLKGDIOTQKGLFFSGRPTPAVYHNLKRESEMEGIPQGEIWMHDIWIKT 651
 Db 594 TLILSGTNLKGDIOTQKGLFFSGRPTPAVYHNLKRESEMEGIPQGEIWMHDIWIKT 653
 Qy 652 FRAENFOIKGSAVSRNVSIEGNTVYNNANATGVPVNOQNTICRSBDWTGLITCOK 711
 Db 654 FRAENFOIKGSAVSRNVSIEGNTVYNNANATGVPVNOQNTICRSBDWTGLITCOK 713
 Qy 712 VDLTQTKYNSIPETQINGINLTDNATVAVKGLANGVTLNHSQFTLSNNAIOGIN 771
 Db 714 VDLTQTKYNSIPETQINGINLTDNATVAVKGLANGVTLNHSQFTLSNNAIOGIN 773
 Qy 772 IRLSDNSTAVDANANGVHLTDSAPSLKNSHFSHQIOGDKGTVLTENATWMPST 831
 Db 774 IRLSDNSTAVDANANGVHLTDSAPSLKNSHFSHQIOGDKGTVLTENATWMPST 833
 Qy 832 TLONLTLNNSVTLSAAYSALSINAPRRRSLEFETTPPSASHPNTLTUNGKLSGGCT 889
 Db 834 TLONLTLNNSVTLSAAYSALSINAPRRRSLEFETTPPSASHPNTLTUNGKLSGGCT 893
 Qy 890 FQFTSSLFGYKSDKLTLSNDAEGDYLLSVRNTGKEPBTLPQTLTVESKDNQPLSDKLTFT 949
 Db 894 FQFTSSLFGYKSDKLTLSNDAEGDYLLSVRNTGKEPBTLPQTLTVESKDNQPLSDKLTFT 953
 Qy 950 LENDVDAGALRYKLVKNDGEPRLNPIKEGSLHNDLVRAEQERTLEAKQVEQTACTOT 1009
 Db 954 LENDVDAGALRYKLVKNDGEPRLNPIKEGSLHNDLVRAEQERTLEAKQVEQTACTOT 1013
 Qy 1010 GEPKRSRRAARAAPDPLPDOSLNLALBAKAEILTAETOKSAKTKKYSKRAV--FSD 1067
 Db 1014 SKARVR--RAVSDPLPAOSLNLALBAKAEILTAETOKSAKTKKYSKRAARFSD 1067
 Qy 1068 PLDDQSLPALBAALVYIDAPQSEKRLAQBAAEK--QKQKDIISYNSALSLSATVN 1126
 Db 1068 TLDPQ--LLOALAEVYIDAQQVKEPQTOBEERKQKCKE--LSRYSNALSLSATVN 1124
 Qy 1127 SMLSVDDELDRFLVDOAASAVMTNIAODKRRYDSDAPRAYQOQKNTLRQIGYOKALNGR 1186
 Db 1125 SMLSVDDELDRFLVDOAASAVMTNIAODKRRYDSDAPRAY--QOKTLRQIGYOKALDNGR 1183

QY 1187 IGAVFHSRSDNTFDEOVKNHATLTMMGSGFAOYOMGDLQPSVNVNGISASKVAEBSRK 1246
 DB 1184 IGAVFHSRSDNTFDEOVKNHATLTMMGSGFAOYOMGDLQPSVNVNGISASKVAEBSRK 1243
 QY 1247 IHRKAINGVNASYOFRLGOLGIPYGVNRETERENYSEBEVAKTPSIAFNRVYAGI 1306
 DB 1244 IHRKAINGVNASYOFRLGOLGIPYGVNRETERENYSEBEVAKTPSIAFNRVYAGI 1303
 QY 1307 RVDTFTPTDNISIKPYFVNVYDVSNANVQTTVNLTLQOPGRVYKVEVGLKAEILHF 1366
 DB 1304 RVDTFTPTDNISIKPYFVNVYDVSNANVQTTVNLTLQOPGRVYKVEVGLKAEILHF 1363
 QY 1367 QISAFISKSQSGOLGKQOONVGVKLGYNM 1394
 DB 1364 QISAFISKSQSGOLGKQOONVGVKLGYNM 1391

RESULT 14 PCT-US03-05226-11

Sequence 11, Application PC/TUS0305226
 GENERAL INFORMATION:
 APPLICANT: St. Geme, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RFT/DCF/DHR
 CURRENT APPLICATION NUMBER: PCT/US03/05226
 PRIOR FILING DATE: 2003-02-18
 PRIOR APPLICATION NUMBER: US 08/296,791
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 11
 LENGTH: 1391
 TYPE: PRT
 ORGANISM: Haemophilus influenzae
 PCT-US03-05226-11

Query Match 72.8%; Score 5297.5; DB 1; Length 1391;
 Best Local Similarity 74.7%; Pred. No. 0;

Matches 1052; Conservative 109; Mismatches 216; Indels 31; Gaps 15;

QY 1 MKTIVRLNLTACISIGISOMAGHTYFGIDYQYRDPFAENKGFVAGNOMIKYNNQ 60
 DB 1 MKTIVRLNLTACISIGISOMAGHTYFGIDYQYRDPFAENKGFVAGNOMIKYNNQ 60
 QY 61 GQAVGTSMTKAPMIDFVSVRNGVALVENGYIVSAHVNGYTDVDFGABENPDQHRFT 120
 DB 61 GEMIGTMKGVPMEDLSMVVGGYSTLISBQHLSTVAHVNGYDVAVDPMGEBNPDQHRFX 120
 QY 121 YKIVRNNYKKNLHPREDYHNPRLHKEFTLEAPLIDMTSNMNGSTYSDRKTPERYRIG 180
 DB 121 YKIVRNNYKKNLHPREDYHNPRLHKEFTLEAPLIDMTSNMNGSTYSDRKTPERYRIG 178
 QY 121 YKIVRNNYKKNLHPREDYHNPRLHKEFTLEAPLIDMTSNMNGSTYSDRKTPERYRIG 180
 DB 121 YKIVRNNYKKNLHPREDYHNPRLHKEFTLEAPLIDMTSNMNGSTYSDRKTPERYRIG 178
 QY 181 SGQOFNRNDQK--GDQVAGAYHYLTAGNTNMQGANGSVYLGQVAKAGEPPLING 238
 DB 179 SGQOFNRNDQK--GDQVAGAYHYLTAGNTNMQGANGSVYLGQVAKAGEPPLING 237
 QY 239 SKGDSGPMFTYDAEKOKMLINGILREBNPEGEKNSFOUVKSYDFEIERDLATSLYT 298
 DB 238 SFEDSGSPMFTYDAEKOKMLINGILREBNPEGEKNSFOUVKSYDFEIERDLATSLYT 297
 QY 299 RAGNGYTTISGNDG-----QSSITQSGGISSEKITTANMGL-PLKAKDKVHNPYRDP 352
 DB 298 TNAEYFNGSDHNGRVALTKSTLPERKAIOBERVLSLJNDSQJHARDKNDSDSPSYKGP 357
 QY 353 NISYPLNNGETLYVNDOKOGSLIFASDINGAGAGLYFGNFTVSPNSNO-TWOGAGHYV 411
 DB 358 NISYPLNNGETLYVNDOKOGSLIFASDINGAGAGLYFGNFTVSPNSNO-TWOGAGHYV 417
 QY 412 SENSITVTKVNGVEHRLSKIGTLIVQAKENGKSGISVQKXILEQADDOGNKQAE 471
 DB 418 GERSITVQVHNPEGRRLSKIGTLIVQAKENGKSGISVQKXILEQADDOGNKQAE 477

QY 472 SETIGISRGTVQNDQKQPTDKEFYFGFGGLDINGHSLTPKRLONTDEGAMTVNNT 531
 DB 478 KKVIGVSGRAVQVQNSADQVDPNNIYFGFGGLDINGHSLTPKRLONTDEGAMTVNNT 537
 QY 532 TQANVITTNESIVLPHNGNNINKLDYRKEILANMGPEDTKNGHNRLLYKPTEDR 591
 DB 538 TQANVITTNESIVLPHNGNNINKLDYRKEILANMGPEDTKNGHNRLLYKPTEDR 593
 QY 592 TLLSGGTNKKDITQTKGKLPFGSGRPTPHAVNHLKMSMEGIPQGEIWMHDIMNT 651
 DB 594 HLLSGGTNKKDITQTKGKLPFGSGRPTPHAVNHLKMSMEGIPQGEIWMHDIMNT 653
 QY 652 FPAENFOLKGSAAVSRVSSIBENMTVSNANATFVGVPVQONTICTRSDMTGLTTCQK 711
 DB 654 FPAENFOLKGSAAVSRVSSIBENMTVSNANATFVGVPVQONTICTRSDMTGLTTCQK 713
 QY 712 VDLDTKIVNSIEPKQINGSTNLTDNAITAVKGLATNGNTLTNHSOFTLSNNAVQGN 771
 DB 714 VDLDTKIVNSIEPKQINGSTNLTDNAITAVKGLATNGNTLTNHSOFTLSNNAVQGN 773
 QY 772 IRLSDNSTATVFNANINGVHLTDSAPSLKNSHFSHQIGDKGTVTLLENATWTMPSPT 831
 DB 774 IRLSDNSTATVFNANINGVHLTDSAPSLKNSHFSHQIGDKGTVTLLENATWTMPSPT 833
 QY 832 TLQNTLTANSTTLNLSAISASSNTP--RRRSLETTTPTSAHRPNTLTONGKLSGQGT 889
 DB 834 TLQNTLTANSTTLNLSAISASSNTP--RRRSLETTTPTSAHRPNTLTONGKLSGQGT 893
 QY 890 POFITSLRGYKSKDKLNSDMDGDTLISVNNYKXEPRETLFQTLTVESKINQPLSDKMT 949
 DB 894 POFITSLRGYKSKDKLNSDMDGDTLISVNNYKXEPRETLFQTLTVESKINQPLSDKMT 953
 QY 950 LENDVDAGALRYKLVKNDGSEFLNFIKEOELHNDLVRABQAEKRLTAKOYEPTAKQT 1009
 DB 954 LENDVDAGALRYKLVKNDGSEFLNFIKEOELHNDLVRABQAEKRLTAKOYEPTAKQT 1013
 QY 1010 GEPKYSRRAPAAPDTLIPQOSLNLNLEAKQELTABTQSAKTKKVSRAV--FSD 1067
 DB 1014 GEPKYSRRAPAAPDTLIPQOSLNLNLEAKQELTABTQSAKTKKVSRAV--FSD 1067
 QY 1068 PLTQOSLFALEALAEVLAPOQSEKDLAQEAEK--QKQKDLISRSNSALSLSATVN 1126
 DB 1066 PLTQOSLFALEALAEVLAPOQSEKDLAQEAEK--QKQKDLISRSNSALSLSATVN 1124
 QY 1127 SMLSYQDELRLFVDDQASAVWTNIAQDKARYDSDAFRAVQOQKTNLQIGYOKALANGR 1186
 DB 1125 SMLSYQDELRLFVDDQASAVWTNIAQDKARYDSDAFRAV--QKQKTNLQIGYOKALANGR 1183
 QY 1187 IGAVFHSRSDNTFDEOVKNHATLTMMGSGFAOYOMGDLQPSVNVNGISASKVAEBSRK 1246
 DB 1184 IGAVFHSRSDNTFDEOVKNHATLTMMGSGFAOYOMGDLQPSVNVNGISASKVAEBSRK 1243
 QY 1247 IHRKAINGVNASYOFRLGOLGIPYGVNRETERENYSEBEVAKTPSIAFNRVYAGI 1306
 DB 1244 IHRKAINGVNASYOFRLGOLGIPYGVNRETERENYSEBEVAKTPSIAFNRVYAGI 1303
 QY 1307 RVDTFTPTDNISIKPYFVNVYDVSNANVQTTVNLTLQOPGRVYKVEVGLKAEILHF 1366
 DB 1304 RVDTFTPTDNISIKPYFVNVYDVSNANVQTTVNLTLQOPGRVYKVEVGLKAEILHF 1363
 QY 1367 QISAFISKSQSGOLGKQOONVGVKLGYNM 1394
 DB 1364 QISAFISKSQSGOLGKQOONVGVKLGYNM 1391

RESULT 15

US-10-687-046-11
 Sequence 11, Application US/10687046

GENERAL INFORMATION:
 APPLICANT: St. Geme, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RFT/DCF/DHR

CURRENT APPLICATION NUMBER: US/10/687,046
 CURRENT FILING DATE: 2003-10-15
 PRIOR APPLICATION NUMBER: US/10/080,505
 PRIOR FILING DATE: 2002-02-22
 PRIOR APPLICATION NUMBER: US 08/296,791
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: Patent version 3.1
 SEQ ID NO 11
 LENGTH: 1391
 TYPE: PRT
 ORGANISM: Haemophilus influenzae
 US-10-687-046-11

Query Match 72.8%; Score 5297.5; DB 31; Length 1391;
 Best Local Similarity 74.7%; Pred. No. 0;
 Matches 1052; Conservative 109; Mismatches 216; Indels 31; Gaps 15;

QY 1 MKCTYFRNLNLTACISLGVSCAWAGHTYFIDYDYDDFAENKGTGAGQNIKYNNK 60
 DB 1 MKCTYFRNLNLTACISLGVSCAWAGHTYFIDYDYDDFAENKGTGAGQNIKYNNK 60
 QY 61 GQVGTSTKAPMIDPSVSRNGVALVENQYIVSAHNGYTDVDCGAEKNNPDHPT 120
 DB 61 GEMGTAKGVMPIDSSVYRGYSTILSGLSVANVGIDVADGAGBNPDHPT 120
 QY 121 YKTVKNNYKKNLHPEYEDDYHNLKFEVTEAPIDMTSNMNGSTYSEKTYPERVAG 180
 DB 121 YKTVKNNYKKNLHPEYEDDYHNLKFEVTEAPIDMTSNMNGSTYSEKTYPERVAG 180
 QY 121 YKTVKNNYKKNLHPEYEDDYHNLKFEVTEAPIDMTSNMNGSTYSEKTYPERVAG 180
 DB 121 YKTVKNNYKKNLHPEYEDDYHNLKFEVTEAPIDMTSNMNGSTYSEKTYPERVAG 180
 QY 181 SGROFWRDQD--GQVAGAHYITLGTATNORAGNGYISLGGDYRAGGYPLAG 238
 DB 181 SGROFWRDQD--GQVAGAHYITLGTATNORAGNGYISLGGDYRAGGYPLAG 238
 QY 179 SGHQMWDKNNKTIQGLAYG--SWLIGATFEDGPRAGNGLTELNGRYQNNKYGPLTAG 237
 DB 179 SGHQMWDKNNKTIQGLAYG--SWLIGATFEDGPRAGNGLTELNGRYQNNKYGPLTAG 237
 QY 233 SKGDSSEPMFYDAEKOKYLLNGILREGNPFEGKEGPGOLVRSKYPEDEIPEEDLTSY 238
 DB 233 SKGDSSEPMFYDAEKOKYLLNGILREGNPFEGKEGPGOLVRSKYPEDEIPEEDLTSY 238
 QY 238 SFGSSGSMFTYKXKYLNGVLRGNYYAAVNGSYQITRKQYRQGLINDITANFWD 297
 DB 238 SFGSSGSMFTYKXKYLNGVLRGNYYAAVNGSYQITRKQYRQGLINDITANFWD 297
 QY 299 RAGNSTVTSIGDNG----QGSITQKSGIPSEIKITLANNST--PLKEXDKVHNPYDGP 352
 DB 299 RAGNSTVTSIGDNG----QGSITQKSGIPSEIKITLANNST--PLKEXDKVHNPYDGP 352
 QY 238 TNAEYRPNIGSDHNGVATIKSTLPKAIQPERVGLYDNGQLDADNGGSESPYKGP 357
 DB 238 TNAEYRPNIGSDHNGVATIKSTLPKAIQPERVGLYDNGQLDADNGGSESPYKGP 357
 QY 333 NTSYRLNGETLYFMDQKOSLIFASDINQAGGLYEBNGFTYSPNSQ--TWQAGLHV 411
 DB 333 NTSYRLNGETLYFMDQKOSLIFASDINQAGGLYEBNGFTYSPNSQ--TWQAGLHV 411
 QY 358 NWSPALHGRSLYFEDQGTGTLTENNINQAGGLYEBNGFTYSPNSQ--TWQAGLHV 417
 DB 358 NWSPALHGRSLYFEDQGTGTLTENNINQAGGLYEBNGFTYSPNSQ--TWQAGLHV 417
 QY 412 SENSIVTWTKNVGEHRLSLKIGKTLFVQKNGKNGSISVGDGVIIHQADDOGNKQAF 471
 DB 412 SENSIVTWTKNVGEHRLSLKIGKTLFVQKNGKNGSISVGDGVIIHQADDOGNKQAF 471
 QY 418 GBSVTEWQVHNPEBRLSKIGLTLVYNKGRNLGSLVANGVLVLDQADESGOKAF 477
 DB 418 GBSVTEWQVHNPEBRLSKIGLTLVYNKGRNLGSLVANGVLVLDQADESGOKAF 477
 QY 472 SEIGVSGRGTQVNDKQPDPTDFYFGFRGRLDLNGHSLTFKRIQNTDEGAMIVHNT 531
 DB 472 SEIGVSGRGTQVNDKQPDPTDFYFGFRGRLDLNGHSLTFKRIQNTDEGAMIVHNT 531
 QY 478 KEVGVSGRATVQNSADQVDPNNIYEGFGRGLDLNGHSLTFKRIQNTDEGAMIVHNT 537
 DB 478 KEVGVSGRATVQNSADQVDPNNIYEGFGRGLDLNGHSLTFKRIQNTDEGAMIVHNT 537
 QY 532 TOANVTITNGESIVLPNGNNINLQVREIAYNGRGETDKKNGHRLNLTLYKFTEDR 591
 DB 532 TOANVTITNGESIVLPNGNNINLQVREIAYNGRGETDKKNGHRLNLTLYKFTEDR 591
 QY 538 SCTANITITGNATI----NSDSKQLTNKKQALFNGPGEODKATNGRLNVAQPYAKEN 593
 DB 538 SCTANITITGNATI----NSDSKQLTNKKQALFNGPGEODKATNGRLNVAQPYAKEN 593
 QY 592 TLLSGGTNLKGDITOTQKGLFSGRPTPHAYHNLKRWSEMEGIPQGIWVDHDMINT 651
 DB 592 TLLSGGTNLKGDITOTQKGLFSGRPTPHAYHNLKRWSEMEGIPQGIWVDHDMINT 651
 QY 594 TLLSGGTNLKGDITOTQKGLFSGRPTPHAYHNLKRWSEMEGIPQGIWVDHDMINT 653
 DB 594 TLLSGGTNLKGDITOTQKGLFSGRPTPHAYHNLKRWSEMEGIPQGIWVDHDMINT 653
 QY 652 FRANKFOIKGGSVSRVSSIRGNVTSSNNAATEGVVPOQNTICTRSDWTGLTCTOK 711
 DB 652 FRANKFOIKGGSVSRVSSIRGNVTSSNNAATEGVVPOQNTICTRSDWTGLTCTOK 711
 QY 654 FRANKFOIKGGSVSRVSSIRGNVTSSNNAATEGVVPOQNTICTRSDWTGLTCTOK 713
 DB 654 FRANKFOIKGGSVSRVSSIRGNVTSSNNAATEGVVPOQNTICTRSDWTGLTCTOK 713
 QY 712 VDLDTFVINSIPKQVINGSINTLQNAATNVGLAKMGNVTLTMSQFTLSNNAQTQEN 771
 DB 712 VDLDTFVINSIPKQVINGSINTLQNAATNVGLAKMGNVTLTMSQFTLSNNAQTQEN 771
 QY 714 VDLDTFVINSIPKQVINGSINTLQNAATNVGLAKMGNVTLTMSQFTLSNNAQTQEN 773
 DB 714 VDLDTFVINSIPKQVINGSINTLQNAATNVGLAKMGNVTLTMSQFTLSNNAQTQEN 773
 QY 772 TRSDMSJATVDNANLNGNVHLTDSAQFSLKNSHSHQIQGDKGTVTLLENATWTMPSDT 831
 DB 772 TRSDMSJATVDNANLNGNVHLTDSAQFSLKNSHSHQIQGDKGTVTLLENATWTMPSDT 831

DB 774 IKLSNHNATVQNNANLNGNVHLTDSAQFSLKNSHSHQIQGDKGTVTLLENATWTMPSDT 833
 QY 832 TLQNLTLNNTITLNSAYSSANNP--RRSLTETPTTSAHNPNTLYNGLSQGG 889
 DB 832 TLQNLTLNNTITLNSAYSSANNP--RRSLTETPTTSAHNPNTLYNGLSQGG 889
 QY 834 TLQNLTLNNTITLNSAYSSANNP--RRSLTETPTTSAHNPNTLYNGLSQGG 893
 DB 834 TLQNLTLNNTITLNSAYSSANNP--RRSLTETPTTSAHNPNTLYNGLSQGG 893
 QY 890 FQFTSLFGYKSDKUKLSNDEGDYLLSVYNTKGPETLEBTLTVESKQNPISDKLFT 949
 DB 890 FQFTSLFGYKSDKUKLSNDEGDYLLSVYNTKGPETLEBTLTVESKQNPISDKLFT 949
 QY 894 FQFTSLFGYKSDKUKLSNDEGDYLLSVYNTKGPETLEBTLTVESKQNPISDKLFT 953
 DB 894 FQFTSLFGYKSDKUKLSNDEGDYLLSVYNTKGPETLEBTLTVESKQNPISDKLFT 953
 QY 950 LENDHYDAALRYKLVKNDGEPFLNPIKXQELNDIVRAEQAERTLAKQVEPTAQT 1009
 DB 950 LENDHYDAALRYKLVKNDGEPFLNPIKXQELNDIVRAEQAERTLAKQVEPTAQT 1009
 QY 954 LENDHYDAALRYKLVKNDGEPFLNPIKXQELNDIVRAEQAERTLAKQVEPTAQT 1013
 DB 954 LENDHYDAALRYKLVKNDGEPFLNPIKXQELNDIVRAEQAERTLAKQVEPTAQT 1013
 QY 1010 GEPVRSRAAANPPDLPDQSLNALAKQAELTETQKSAKTKVRSKAV--FSD 1067
 DB 1010 GEPVRSRAAANPPDLPDQSLNALAKQAELTETQKSAKTKVRSKAV--FSD 1067
 QY 1014 SKARVSR--RAVSPDPAQSLKALBAKQ--LTETQTSVA--KVRKRAAREFSD 1067
 DB 1014 SKARVSR--RAVSPDPAQSLKALBAKQ--LTETQTSVA--KVRKRAAREFSD 1067
 QY 1068 PLDDQSLFALMALVYIDAPQSEKDLAQBARK--QRKQDLISRYNSALSLSLFTVA 1126
 DB 1068 PLDDQSLFALMALVYIDAPQSEKDLAQBARK--QRKQDLISRYNSALSLSLFTVA 1126
 QY 1127 SMLVODELDELFDVQASAVNTNIAQDKRYSDAFRAYQOQKTKLQIGVQALANG 1186
 DB 1127 SMLVODELDELFDVQASAVNTNIAQDKRYSDAFRAYQOQKTKLQIGVQALANG 1186
 QY 1125 SMLVODELDELFDVQASAVNTNIAQDKRYSDAFRAYQOQKTKLQIGVQALANG 1183
 DB 1125 SMLVODELDELFDVQASAVNTNIAQDKRYSDAFRAYQOQKTKLQIGVQALANG 1183
 QY 1187 IGAVFSHRSQDTEDEQVXNATLTVAGGFAQYQMGDLQGVVNGTGISAKMAEBSRK 1246
 DB 1187 IGAVFSHRSQDTEDEQVXNATLTVAGGFAQYQMGDLQGVVNGTGISAKMAEBSRK 1246
 QY 1184 IGAVFSHRSQDTEDEQVXNATLTVAGGFAQYQMGDLQGVVNGTGISAKMAEBSRK 1243
 DB 1184 IGAVFSHRSQDTEDEQVXNATLTVAGGFAQYQMGDLQGVVNGTGISAKMAEBSRK 1243
 QY 1247 IHRKALNGVANSYQFLGQLGIPYGVNRYFIERENTQSEBYVKTPLAENRYNAGI 1306
 DB 1247 IHRKALNGVANSYQFLGQLGIPYGVNRYFIERENTQSEBYVKTPLAENRYNAGI 1306
 QY 1244 IHRKALNGVANSYQFLGQLGIPYGVNRYFIERENTQSEBYVKTPLAENRYNAGI 1303
 DB 1244 IHRKALNGVANSYQFLGQLGIPYGVNRYFIERENTQSEBYVKTPLAENRYNAGI 1303
 QY 1307 RVDTFTPTDNTSYKPYFVNVYDVSNNVQTTNLTLYQOPRGYXNOKYGLKAEILHF 1366
 DB 1307 RVDTFTPTDNTSYKPYFVNVYDVSNNVQTTNLTLYQOPRGYXNOKYGLKAEILHF 1366
 QY 1304 RVDTFTPTDNTSYKPYFVNVYDVSNNVQTTNLTLYQOPRGYXNOKYGLKAEILHF 1363
 DB 1304 RVDTFTPTDNTSYKPYFVNVYDVSNNVQTTNLTLYQOPRGYXNOKYGLKAEILHF 1363
 QY 1367 QLSAFISKSQGSOLAKQOONVGLGYRW 1394
 DB 1367 QLSAFISKSQGSOLAKQOONVGLGYRW 1394
 QY 1364 QLSAFISKSQGSOLAKQOONVGLGYRW 1391
 DB 1364 QLSAFISKSQGSOLAKQOONVGLGYRW 1391

Search completed: June 23, 2004, 17:37:12
 Job time : 231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw mode.

Run on: June 23, 2004, 17:30:42 ; Search time 21 seconds

[without alignments]
3036.187 Million cell updates/secTitle: US-10-645-655-2
Perfect score: 7273

Sequence: 1 MKKTYFRLNPLTACISLGLIV.....SQSGQLGQGVKGLGKRW 1394

Scoring table:

BL0SUM62

Gapcp 10.0, Gapext 0.5

Searched: 302849 seqs, 45738825 residues

302849

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Pending patents AA New:*
1: /cgn2_6/prodata/1/paa/PCT_NEM_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1609	22.1	1694	1 PCT-US04-06615-4	Sequence 4, Appli
2	1529	21.0	1794	1 PCT-US04-07001-763	Sequence 763, App
3	295.5	4.1	1542	1 PCT-US04-07001-673	Sequence 670, App
4	285.5	3.9	1004	5 US-09-806-709-57	Sequence 57, Appl
5	285.5	3.9	1010	5 US-09-806-709-55	Sequence 55, Appl
6	285.5	3.9	1222	5 US-09-806-709-37	Sequence 37, Appl
7	283.5	3.9	1228	5 US-09-806-709-34	Sequence 34, Appl
8	283.5	3.9	1536	5 US-09-806-709-67	Sequence 67, Appl
9	280	3.8	1477	5 US-09-806-709-71	Sequence 71, Appl
10	277	3.8	977	5 US-09-806-709-53	Sequence 53, Appl
11	277	3.8	983	5 US-09-806-709-51	Sequence 51, Appl
12	275	3.8	1381	1 PCT-US04-07001-693	Sequence 693, Appl
13	275	3.8	1492	1 PCT-US04-07001-672	Sequence 672, App
14	269	3.7	1005	5 US-09-806-709-41	Sequence 41, Appl
15	269	3.7	1011	5 US-09-806-709-39	Sequence 39, Appl
16	246.5	3.4	969	5 US-09-806-709-32	Sequence 32, Appl
17	246.5	3.4	975	5 US-09-806-709-30	Sequence 30, Appl
18	245.5	3.4	1180	5 US-09-806-709-65	Sequence 65, Appl
19	245.5	3.4	1188	5 US-09-806-709-63	Sequence 63, Appl
20	245.5	3.3	1073	5 US-09-806-709-49	Sequence 49, Appl
21	242.5	3.3	1079	5 US-09-806-709-47	Sequence 47, Appl
22	242.5	3.3	1290	5 US-10-786-892-138	Sequence 138, Appl
23	239	3.3	915	5 US-09-806-709-35	Sequence 35, Appl
24	238	3.3	10203	6 US-10-724-972A-4098	Sequence 4098, Ap
25	237	3.3	1035	5 US-09-806-709-63	Sequence 69, Appl
26	235	3.2	1531	1 PCT-US03-32645-44	Sequence 44, Appl

27	232	3.2	1336	5 US-09-806-709-73	Sequence 73, Appl
28	229	3.1	1395	5 US-09-806-709-45	Sequence 45, Appl
29	229	3.1	1101	5 US-09-806-709-43	Sequence 43, Appl
30	228	3.1	1220	5 US-09-806-709-28	Sequence 28, Appl
31	228	3.1	1226	5 US-09-806-709-26	Sequence 26, Appl
32	221	3.0	906	1 PCT-US04-12717-56	Sequence 56, Appl
33	221	3.0	906	6 US-10-831-070-56	Sequence 56, Appl
34	218.5	3.0	915	1 PCT-US03-31057-3	Sequence 3, Appl1
35	218	3.0	992	5 US-09-806-709-61	Sequence 61, Appl
36	218	3.0	998	5 US-09-806-709-59	Sequence 59, Appl
37	212	2.9	3636	6 US-10-724-972A-5942	Sequence 5942, Ap
38	205.5	2.8	3829	1 PCT-US03-38966-10	Sequence 10, Appl
39	204.5	2.8	574	1 PCT-US03-31057-73	Sequence 73, Appl
40	204.5	2.8	578	1 PCT-US03-31057-141	Sequence 141, App
41	204.5	2.8	579	1 PCT-US03-31057-140	Sequence 140, App
42	204.5	2.8	580	1 PCT-US03-31057-139	Sequence 139, App
43	204.5	2.8	581	1 PCT-US03-31057-138	Sequence 138, App
44	204.5	2.8	582	1 PCT-US03-31057-137	Sequence 137, App
45	204.5	2.8	583	1 PCT-US03-31057-136	Sequence 136, App

ALIGNMENTS

RESULT 1
PCT-US04-06615-4
Sequence 4, Application PC/TUS0406615
GENERAL INFORMATION:
APPLICANT: New England Medical Center Hospitals, Inc.
APPLICANT: Plant, Andrew G
TITLE OF INVENTION: Treatment of IgM1 Deposition Diseases
FILE REFERENCE: 28154/2068
CURRENT APPLICATION NUMBER: PCT/US04/06615
CURRENT FILING DATE: 2004-03-09
PRIOR APPLICATION NUMBER: US 60/453055
PRIOR FILING DATE: 2003-03-07
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patent version 3.2
SEQ ID NO 4
LENGTH: 1694
TYPE: PRT
ORGANISM: Haemophilus influenzae
PCT-US04-06615-4

Query Match 22.1%; Score 1609; DB 1; Length 1694;
Best Local Similarity 26.5%; Pred. No. 1.5e-78;
Matches 479; Conservative 251; Mismatches 536; Indels 542; Gaps 52;

QY	6	PLNPLTACISLGLIVSQAMAGHTTGGIDYOTRDPENKKGFTVGAONITVYKKGQVVG	65
DB	6	PLNPLTACISLGLIVSQAMAGHTTGGIDYOTRDPENKKGFTVGAONITVYKKGQVVG	65
QY	66	TSMTKA-PMIDFSVVS-RNGVALVENQYVSVAH-NNGYTDVDFG-AEGN-----NPDQH	117
DB	66	NVLPMGSLPMIDFSVVDVDRKATLILINPQYVGVKHSNGVSELFGLNLCNNMNGNAASH	125
QY	118	RFT-----YKIVKRNXY-----KKONLHPEYEDYHNPRLHKEVTEAPID-MTS	160
DB	126	RVSSEERKPYSEKNEYETKNGKAYATEDQKRRBYVWPLDKEVTEAPIDASTA	185
QY	161	NNNGSTYSIDRTKPEPVIRGSGRQFMRNDKDKQ-----VAGAHYLT	203
DB	186	SSDAGTYNDOKKPAVRLGSGSQF-----LYKKDNYSHLNNHEVGNKIKLGAAYTGG	242
QY	204	TGNTHNORAGAGNGSYLDC-----GVRAGSEYGLPIAGSKDSSGSPFIYDAEKQW	257
DB	243	INGTYKKNHNNGLIGFNSKEHSPKXILSOPPLINAVLADSSPLFIYDAEKQW	302
QY	258	LINGTLBGNPFEGKE-NGFQIVKSYFDEIFRDLHTSLVTRANGVTTISGNDGGS	316
DB	303	LFLGSDYFMAGYKKKQWQBNVITYRPEPAKTVLDKDTAGSL-----TGS	345

QY 317 ITKSGIPSEIKITLANMSLPLKXKCVHNPBYD--GPVYSPRLANGETLYMNOQOSL 375
 DB 346 NTQYNNMPCCTKSVISNCS-----BSJLVNDIPSSQUTDSKNNKXSVTL--KSGSL 397
 QY 376 IPASDINOAGAGLYEGNFTVSPNSNQ-TWQAG-IHVSSTVYMKVNGVEHDLKSCXG 434
 DB 398 TLNNIDOGAGGLFPGDYEYKGTSDSTTMKAGVSAVDGTYTMKYNPKSDRLAKG 457
 QY 435 GTLHVQAKGKNKGS-SVQDGVILIEQADDOCKQAFSEGLVSGGVQVQNDKQFDTD 494
 DB 458 GTLIVEGKGNKSLKVDGTVILKQADANNKVFQVQVIGVSGSTVYVNDKQYDEN 517
 QY 495 KEYFGFGRLJLNGHSLTFKRIQNTDEGAMVYNNITQKANTINGESIVLYR----- 549
 DB 518 SIYGFRRGRLDANNGNLLFHHIRNIDGRLVNNHNSKSTVITIGSSLTIDNTITPY 577
 QY 550 ----- 549
 DB 578 NIDAPEDNPVAFRITKDGQLYNLNENTYVALKKGASTRESELPRNGSGENMLYMK 637
 QY 550 -----GNINKLDYRKAIYNGMFEETDKAKKNGRLNLYKFTTEDRTLLSGTNLKG 603
 DB 638 TSDAKANVYNNHINERNMNGFYGB--EEGKNGNINLVTFKSKSEQRNFLLTGQTYLNG 696
 QY 604 DITQTKKLFSGRPTPHAVY--HLKXSEMEG-POGELVMDHMINRTKAKNPGIKG 661
 DB 697 DLKFKGSLTFSGRPTPHAPADIAQISTYKQDHFENNEVWEDVIMKKNKNTINLVN 756
 QY 662 GSAYVS-RVYSSIEGNMTVSNANATFGVVPNOQNTICTSMVGLTTCQKVDLDTKYI 720
 DB 757 MVLVSGRVANITSNITASDNKXHIQY--KAGDVCVMSDVTGVITCTTDKSLD-KAL 813
 QY 721 NSIPKTVINGSLNLDNATVANGIAKLKQVNTLTNSQSTLSNNAQTQIRLSDNSTA 780
 DB 814 KSFNATVSGVNLGSGANF-VLGKANLFGISGTGNSQ----- 851
 QY 751 TVNNALMGNVHLCPSAOFSLKNSHSHQIQDQKGTVLENAVTMPDITLQNTLUN 840
 DB 852 -----VNL-----MNSHMLGDSNVQNLKX 875
 QY 841 STTLMSAYSASNNTPRRRSLETETPTLSAHEFNLTJYNGKLSGQSTPOFTSSLFQYX 900
 DB 876 GHILNAQNDARKVT-----YNLTLYN-SLGSNGSFYLLDLSNKO 916
 QY 901 SDKLXSNABGDYILSVRYTGEPEPTLEQTLVSGNDQPLSDKLKTLBNHVDAGL 960
 DB 917 GKVAVTKSATGNFTLVADKTGP--TKNELTLFDA--SNATRNINVLWNTVDLIGAM 973
 QY 961 RYLVKNDGEFRLNPIKEOELHNDLY----- 987
 DB 974 KYKLAVNGRYOLYNF--EVEKNGQVDTMTITTPNNIQADVSPVSNNEELAVETVP 1031
 QY 998 -----RAEQAEETLE-----AKOVEFLAK--TQGE-- 1011
 DB 1032 PPAPATSEETETVAENSQKSEKVEKNEQDATETTAQNGEVAEFAKSVYANTQINEVA 1091
 QY 1012 -----PKYSSRAARAF-----DTL 1028
 DB 1092 QSGSSEETQTEKETAKYKGEBAKVEKOEIOEAPQASSTSPKQKAPKREYSTDK 1151
 QY 1029 PDQSLNLAL-----EAKQELTATQSKA----- 1053
 DB 1152 VEBTQVQAPQOSTVVAEAATSPNSKPAETQPSBKTNAPVPIVVSXKQNTENTDP 1211
 QY 1054 -----KTKYRSKR-----AVSPLLDQ----- 1072
 DB 1212 TEREKTAKEVEKTEKQEPPOVASQSPKQEBETVOPAVLESENVPTVNAHEVQACQOT 1271
 QY 1073 -----SLFAELAVITAPQO----- 1089
 DB 1272 QTSATYSTKQAPAPENSINTQSATLTETAEKSEKFPOTETMASTEDASCHANTVADNSVA 1331
 QY 1090 -----SEKRLAQSE-----AEKOROK----- 1107

DB 1332 NNSBSSDKSRBRKSIQPOETSAEETTAETDTEITLADNSKSKXPNNRSRVRSEPTV 1391
 QY 1108 -----DHSYNSALSLSLA-----TVNSMLSYODELDELFLV--DOAQAVWT 1149
 DB 1392 TNGSRSRSTVALRDTSTNTNAVTSIDMAKQAFVALVNGKQVSQHSOLENNNGQYVWV 1451
 QY 1150 NIAQDRRYSDDAFRAVQOQKTNLRQGVOKALANG-RIGAFVSHSRSDNTEDEQYKMA 1208
 DB 1452 SNTSMNENYSSQRRFSSKSQ-TQSGMQTISNNVQGVFTYVNSNNPD-KASSKN 1509
 QY 1209 TLTMMSGAQY-----QSGDLQGVNVGTGISASMAECSRXIHKALINVNASYQFEL 1264
 DB 1510 TLAQVNFYSKYADNEM--YLGIDIGYKFGQNLKTNHNAFAHHTAQGLTRAKAFNL 1566
 QY 1265 GOLGQZYFYGVNRYPIERENYQSEBEVYKTPSLAFNRVNAGRVDYFTTIDMSYKPY 1324
 DB 1567 GNFGLTPLYGVRSYSSNANFALADRIKONF-SVYTAAPQVDSLYTY-HLSEFVTPIL 1625
 QY 1325 FVNVYVSNNAVQTYNLTVLQOPFGRYQKEVGLKAEILHFOISAFISKSGSQGLKQO 1384
 DB 1626 SARY-DTNGSGKLTNVQYDPAYNVENQOQYNNAGLTKYHVKLSLIGULTYKQAKQK 1684
 QY 1385 NNGVYLG 1392
 DB 1685 TALKLSE 1692

RESULT 2
 PCT-US04-07001-763
 ; Sequence 763, Application PC/TUS0407001
 ; GENERAL INFORMATION:
 ; APPLICANT: BKALETZ et al.
 ; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
 ; FILE REFERENCE: 28335/38815A
 ; CURRENT APPLICATION NUMBER: PCT/US04/07001
 ; PRIOR FILING DATE: 2004-03-05
 ; PRIOR APPLICATION NUMBER: US 6C/453,134
 ; PRIOR FILING DATE: 2003-03-06
 ; NUMBER OF SEQ ID NOS: 771
 ; SOFTWARE: Patceptin version 3.2
 ; SEQ ID NO 763
 ; LENGTH: 1794
 ; TYPE: PRT
 ; ORGANISM: H. influenzae
 PCT-US04-07001-763

Query Match 21.0% Score 1529; DB 1; Length 1794;
 Best Local Similarity 25.4% Pred. No. 3,3e-74;
 Matches 466; Conservative 238; Mismatches 536; Indels 654; Gaps 60;

QY 6 FRLNPLTACISIGIVSOAMAGHTYFGIDYVYRDEFAKKGKFTVGAQNTKYNNKQGLWG 65
 DB 6 FKLNFIALTYAVALPTPEALVANDVDYQIFRDFAKKGKFGVGAINTVEVDNKNLIG 65
 QY 66 TSMTK-APMIDSVYS-RNGVIAVENVQIYVAH-ANGITDVFQ-ASGN-----NPD 115
 DB 66 SALPKDIPMIDSAVDKRLATLVNPQYVGVKVGNGVEJHFNGLNGMNPFGNSI 125
 QY 116 QHRFT-----YKIVRNNY-----KKDLH-----PYE----- 138
 DB 126 QHRVSWEBKRYTYEKNNSESELMKGTQMEKDKQYTSNKKVDPSELYGOLVYEQNQ 185
 QY 139 ---DYHNPPLKHFVTEAPID--WTSNMGSTYSQRTYPERVYIGSGQF----- 185
 DB 186 KREDDYMPRLDKFTVAVAPTEASTISSDAGTYNDQNKYPAFVRLSSGQFYKXGSHYE 245
 QY 186 -----WRNDQKGVYAGAYHILTQGNTHNQRGAGNGYVYG-----GD 224
 DB 246 LILKKNKKEKELIHRMDVGDGMLKLVGNAYTYGLAGPYKVNHTDGLIGFDDSTEDHND 305
 QY 225 VRKAGEYGPLTAGSKGDSGSPMEITYDAEKQMLINGILBEGNPFEGKE--NGFQLVRSY 283

```

Db      306 PKELSRKPLINAVAGDSGSPLFVYDKSKEKMLFLGAYDFWGYKKKSWQENNYKQFQ 365
Qy      284 FDEIFERDLHSLYTRAGNGVYTCISAKNGQGSITQSGSIPSEITKITLANNSLFLKQK 343
Db      366 AENILKCK--SAGLKGNTQY--NMTSGKNTSLISGSESLVDYD----- 408
Qy      344 VHNPRYDGNISYSPILNNGETLYFMQKQSLIFASDINGAGGAYFEGNTYSPNS-NQ 402
Db      409 -----MKNLHGQNVF--ESSGNLILNNMIDQAGGLFFEGDYEVAGTSEMT 454
Qy      403 TWGAGIYVSENSYVYKRVNGVEHRLSKIGKGLIYVQAKENGKSIISVQKYLLEQQA 462
Db      455 TWKAGISYAEKTYKWKVYVNPQFRLAKIGKGLIYVGNQKSKSLAVGQYVYVAKQ 514
Qy      463 DQGNKQAFSEGLVSGRGVQVQANDXQFDTDKYFRRGGIIDLNGSLTYKRIQNTT 522
Db      515 TT--GGHAFASVIGVSGRSTYVANDQVDPNISYFGFRGRLDANGMLFFEHINID 572
Qy      523 GAMIYVHNTQANVTITG--NESIV-----LPYGNL----- 553
Db      573 GARLVYHNTYNSNITITGAGLITPSCVITTYTALTADDNVYVPSIPRGKDLFPM 632
Qy      554 -----MKLDYKKEI-----AYNGMFG 569
Db      633 CYKXYALKQSGSPTAEMP CYSEKSDANWEMECQNDQKXAMVYINRRMNGFNGYFG 692
Qy      570 -ETDKKKNKGNLNIYKPTEDKRLISGQNLKGDITQTKKGLFSSRPPIYVYHLM- 627
Db      693 EAKTADQNGKLVNTFSGKSDQRFLLTGGINLNGELKYEKKTFLSLGRPIYHARDIANI 752
Qy      628 -----KMSMEGIPQGEIYMDHMINRTFKAEFQIKGSAVYS-RVYSISLGWYTS 680
Db      753 SSTEKDKHFAE--ANEVYEDDMIKRTIKANINVTYNNATLVSGRNBSITSITLS 807
Qy      681 NYANATPGVVPNQNTLCTRSQDTGLITQXVDLITDYKINSIPQOYNGSINLTDNATA 740
Db      808 NYAKYHIGY--KAGDYCVASDYGTWYTCAN-DLSTKALNSFNPTNLRGNVLTESA-- 862
Qy      741 NYKGLAKLNGVTLNHSOPTLSNNAIQIGIRLSNSTRATYDANAKNGVHLTSDAQFS 800
Db      863 -----NFTLG-----KANLEGTINSTENSCYN 884
Qy      801 LKNSHFSQIQDGKGTVLENAATYMPSDTLQNTLNNSTITLINSAYSASSNNPRRR 860
Db      885 LK-----ENSHVILNGSDVQICJLNGHILNNNSDRIK----- 919
Qy      861 SLETTPTLSAHRPNTLTVNGKLSGQSTFQTSLSFGYKSDKGLSNDAGDYIISYVN 920
Db      920 -----ETKYHTLNIS-NLSGNSFYVWDFTKQNGKVVYTKSAKGFILQYAN 967
Qy      921 TKKEPTELEQULIVESKQNPJSDKAKFTLENHDVAGLAKYVNDGFRJANFYKQ 980
Db      968 KXGEBPNH-NELTLEFDS-SNATERSGANVSLANGVDRGAMSYTLKENSGRYYJANNEVER 1025
Qy      981 -----ELHNDLYR----- 988
Db      1026 RQVYDTPSITANNQADVPYSNNHEETARVEADIPPLAPAPANGSAMANEOPETRP 1085
Qy      989 ARQARLLE----- 997
Db      1086 AETVQPTMEDTNTTHSGSEPOADTTQADPNSESVPSETEKVAENSPOSETYAKQRO 1145
Qy      998 -----AKQVEPT--AKTOT----- 1009
Db      1146 KATETTAQNDVAKKAPTEANTQVIELAONGSETEETOAEARAROSEINSTEETVED 1205
Qy      1010 -----GEPKVSRR-----JABAAF----- 1024
Db      1206 DPTISEPKSRPSSISSNNINLAGTDTAKYETKTOEAPQVAFASPKQEBBPBAKQ 1265
Qy      1025 -----PDTLPDQ-----SLANALEA----- 1039
Db      1266 QHQPVTQSQAKPEITTTQAEAPARENVSYVANNYKQAPAKPTVAKEETTASNEQKET 1325

```

```

Qy      1040 -----KQABLPAETQYSKANT-----KKVYSKAV 1064
Db      1326 AQVPANPKTAAENKAEQSTETTEDENIHOEPAITAVASTVEPATTPRIKENVKTE 1385
Qy      1065 FSDPILQOSFALE-----AALBYI----- 1084
Db      1386 AQQVTEITTVSITENPVKNPENTTPATQSTVNSEAVQSETATTEAVWSQKTSABET 1445
Qy      1085 -----DAPQ--QSEKDR-----LAQBEAK-----QKQKDL 1109
Db      1446 TYASQETTYDNGSGTTPQPSRRRBRQAQNSYEPPELITENNAENPQGDVVALQVLRDL 1505
Qy      1110 ISRRNSGALSLSA-----TVNSMLSYODELRLFY-DQAQSAVYTNIAQDKRYDSDAF 1163
Db      1506 TETVYNAVITSDAKAKQFVALNVKRAVSQGISQLENNNGQYVWVWNTSMKENYSSQY 1565
Qy      1164 RAYQOQKTNARQVQKAL-ANGRGAVFHSNSDNTPEQVKNHALUTMMSGFAY-- 1219
Db      1566 RHF--SSKSAQTQGMQDTISSNVQLGVTYVNSNNPD-KASSKNTLAQANLYSKTYMD 1623
Qy      1220 -QWEDLOFVNVGTGTSASKAEEQSRIKRAKINGVNVAAYOFRLQGLIOYFGVARY 1278
Db      1624 NFM---YIADLYGTFQFQSNLOTHNAKPAKHTAQFGSLAKAENLGNFAVRYGKYS 1660
Qy      1279 FIERENTQSEEVKPTPSLAFNRNAGIRVDYTFPTDNISVKKYFFVNYVDYSNANVT 1338
Db      1681 YLSNANPALAKODIKVNPISVTAFAQVDLSYTY-HLGEFSITPLISARY-DANQSGSKI 1738
Qy      1339 TVNLTVLQCPFRGYWQKVGKALILHFOISAFISQSGSLGQGVQVYKXGY 1392
Db      1739 NVDRYDPAVNVKQOQYNAKGLKTHNVKLSLIGLITRAKQAEKQAEVXLSF 1792

```

```

RESULT 3
PCT-US04-07001-670
; Sequence 670, Application PC/US0407001
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: PCT/US04/07001
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: Patentia version 3.2
; SEQ ID NO 670
; LENGTH: 1542
; TYPE: PRT
; ORGANISM: H. influenzae
; PCT-US04-07001-670

```

```

Query Match      4.1%; Score 295.5; DB 1; Length 1542;
Best Local Similarity 19.8%; Pred. No. 6,2e-06;
Matches 299; Conservative 202; Mismatches 501; Indels 509; Gaps 74;

Qy      42 ENKGPVTGAKNIKYVNGQVGTSMTRAPMIDFVSVRGVAALVENQYTVSAHVNG 101
Db      219 KNEGVISVNGSISL--AQGKITIDILNPIITYSI-----AAPEEAL-----NLG 264
Qy      102 YTDVDFGABGNPDDHRTTYIKYKRNKYKQNLHFEYDDYHNERLAKFYTEAAPIDMTSN 161
Db      265 -----DIFAKGNINVRATI--RKK-----GKSLASVSXKSGSN 298
Qy      162 MNGSTVSDRTKYBEHVRIGSRQFWRNDQKGD--VAGAVHYLTAGTHNRGAGAGYS 219
Db      299 IILISAKGSAET-----SVTISAQNOQAKGKMTTGQVTLTKGAVYIDLSGKSGSET 351
Qy      220 YLGQDVAKAGEYPLIAGSKDQSGSPMETYDAEKQYMLINGILRGNPF--BKGNGFQ 277
Db      352 YLGQDERGGRNG-----IGLAKTTS-----LEKSGTINVSKEKGR 389

```

```

QY 278 LVKSYFDEIFERDLHTSLYTRAGNYTISGNDGSSITQXSG-----ISSE 326
Db 390 AI-----WMDTALIDGINAKQSDIAKIGVETSGHYLSIDSN 429
QY 327 I-----KILANMSLPLEKXD-KVHNPBYDGNITYSRLNNGEYL----- 365
Db 430 ALVKTKEWLDPDWYITIEAPLSRADTJISSEPIIDGFE-NSPKKADTITILMETISN 486
QY 366 YEMOOK-----QCSLIPASDINOAGGLYFBSNFTVSPNSQWQ 405
Db 489 FLONAKVNTIARUKLTVNSSIISGSSHLIHSE-QQSDGVDIDEDIT-SEGGLTIN 546
QY 406 GAG-IHVSNSYTFKAVNGVEHDLKXIGTLYHAKENKKSISVGGOKYLLEQOQD 464
Db 547 SGGWVDYHKNIT-----LGTFLINTA-----GSSVAHEKG----- 577
QY 465 OGKQAFSEIGLVSGRTVOLN-DDKQFDKFEYFPGRLDINGSLTFKRIQNTD-- 521
Db 578 -GNABRNATQITAGTITVXKDKQ-----FRNNVSIINGSLKFLANQNF 627
QY 522 -----EGMYVNHNTTOA-----NV-TTGNESI-----VLPNGNI 553
Db 628 TKEDEINISGLVITNQTCKADKWHMWSKSYWVWVSLTLNDAKETFEIKEYDSGNS 687
QY 554 NKI-DYKKEIA--YNGMFGEDKQ-KHNGRLNLIXEPT----- 587
Db 688 QULBARRRPAHGHGRTGKTNRINISAKKALFKLKENAATDPKELPITNANITATG 747
QY 588 TEDRTL-----LISGGTINKGDIYTKCKLFFSGRPTPI-----AYNLAN 627
Db 748 SDDSVWFEDIHANLTSRAASINMDSINITGLDFS--ITSNRSNAFEIKKDLINANT 805
QY 628 KMESEBGIPOGFIWDHDMINRTFKANQCKGSSAYV--SRVSSISGNTYSNNANA 685
Db 806 SEFSKQTKDLFENQYTGALNST--RNLTIGKAVTLGGENSSNITGITIAEANY 862
QY 686 TGVVPOQNTICTSDMT-----GLTCCQVUDTDTK-----YINSIP-KIOINSI 732
Db 863 TLQAYAD--NSIKKHKKTILHGNVTSISGSLTSKYEKGDLAVLNAGATKGTINDSL 920
QY 733 NI-----TDNATANY--KGLAKINGNYT-----LJNHSQ-----F 760
Db 921 NITGFTNNGTADINIKGAVNIQGDITNKGKLNITTAQCKQKTIINGNTKKKNLNI 980
QY 761 TLSNNAFOI-----GIRLS-----DNSTAYNANL--N 788
Db 981 TNNNDTEIQIGNISQKESMLTISDKNITKOITIKAGVDEKSSSTASDANLTIXT 1040
QY 789 GNHILDSAQFSLKNSHFSHOIOGDKGTVTLEN-----ATWTFPSDITLQNTL 837
Db 1041 KMLVVEDLNISEFN--KAEITAKGOSDLITGNTNSADGNAKVTINQVYDS---KIS 1094
QY 836 IANSTTILNSANSASSNN-----TPRRSLEFETPTSAEHRFNTLTVNGKL 884
Db 1095 ANDHNVTLSKSVETSGNDTNTGDGSGNNAGLITIAKNEVKNNTIS--AKVTINIASEKL 1152
QY 885 SGO-----GTQFOTSLFGYKSDYKLSND--ABSDYILSVRTGGEPTLEQL 931
Db 1153 TTKADATIAATTGNEVTAKTGDIKGVKSISGNVNIYANGD-TLWVSVSGNAVIT-- 1208
QY 932 TVESKNOPLSDKLTLENDHYTAGA-RYKLYVKNDEGF--RLHNPIDEGELINDVRA 989
Db 1209 -----TADKGLTTOSSSTISGTSVITSSQSDIGANISGNTVSVKATNDLI- 1256
QY 993 EOARTELEAKOVELPAKTQGEPEVTSRRAPAAEPDLPDQSL-----LANLEAK- 1040
Db 1257 TKANSKLEAK-----TSEANVVS--ANGIIIGTISGNTVANTAGSLTITKGA 1304
QY 1041 -----QAEITAE--TOKSAKTKVRSRAVFSPLIDQSLFALEAALE----- 1082
Db 1305 VDAENGATITAESGKLTTOAGSTISSNGQTTITAKDGSIASSIDANANVTINTTGLIT 1364
QY 1083 -----VIDAPQOSEKDLAQBEEKQKOKDLISRYNSALSLS----- 1122

```

```

Db 1365 VQSSINANEGLTYINA-----QDATLNGDASGDRTVNAVNASGSGVTAVTSSVNI 1419
QY 1123 ---ATVNSMLSVOD-----ELDLFVD-----OQASAVNTNIQDREYDSPA 1164
Db 1420 GDLSTINGNITISNKGKNTVYALKAELDYKTIQGVASAEFVIEAKRLLEVKLSI--- 1476
QY 1465 AYQOQKTNLRQIGYOKALANGRIGAVFESHSSDNTFDEQYKHAATLWMSGFAQYQWGD 1224
Db 1477 ---EERETLAKGVSAVFEPEPMALITVTONFEPT---TRPSQVITSEKACFSSGD- 1528
QY 1225 QPGWNYGTGIS 1235
Db 1529 --GAAVCTIVA 1537

RESULT 4
US-09-806-709-57
Sequence 57, Application US/09806709
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
FILE REFERENCE: 1038-1138 MTS
CURRENT APPLICATION NUMBER: US/09/806/709
PRIOR APPLICATION NUMBER: PCT/CA99/00938
PRIOR FILING DATE: 1999-10-07
PRIOR APPLICATION NUMBER: 09/167,568
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 09/206,942
PRIOR FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 57
LENGTH: 1004
TYPE: PRF
ORGANISM: Haemophilus influenzae
US-09-806-709-57

Query Match 3.9%; Score 285.5; DB 5; Length 1004;
Best local similarity 21.7%; Pred. No. 1.2e-07;
Matches 254; Conservative 169; Mismatches 438; Indels 309; Gaps 59;

QY 101 GYTDVDFGABGNPNPDQREFTYKIV--RNN-----YK 130
Db 25 GREGDASNAKXNNHDKPTLNTTVENALKNTFPNITAKKLTIVNSDINTIKGAHLITLXS 84
QY 131 KXNLEPZEDDYHNRFLHKTVTAEADITKTSNMGSTYSRTPPE----- 175
Db 85 KNN-----KXSSVKNKINGNISTTNGNLTLYSSGWDIHKNTITNTGY 126
QY 176 -PRFISGROFPWNRDDKGDQVYA-----GAVHYLTAGNTNQGAGNGSYVGGDVAKGE 230
Db 127 LNTLGGSVAFERKAKNEKRGVSESVYKAGCVITS-----GVQEGFRF--NNVSLKCY 177
QY 231 YGPPIASGKSDSGSPMFTYDAEKOKLINGILLBGNPEKENGFOVLVKSYFDEIFER 290
Db 178 GAGLAFVQKNISSN-----SVRENTI--KNRFDGALNLSGVAVNS--DVSGT 222
QY 291 DLHSLYLRAGNGYITISGNDGOSIT-QKSGIPEIKTITAMKSLPLEKOKVKNRPY 349
Db 223 KWHTRINRYWVNTLINAAGSSSFNLSIDASSISGMDDTYRGL-----NGITF 274
QY 350 DGNITYSRLNNGEITYF-----MDQKGS--LIPASDIN-QCAGGLYEGNFTVSPNS 400
Db 275 NBNFTN--LQGSINAHFKTSVVTPEKPSNATLNGNISTUGGTYVHELN--ASSS 329
QY 401 NOTWQAGIHVNSNTVTVKVGVEHDLKXIGKTLHYQAKG-----ENKGSISVG 452
Db 330 THTTSGALIN-SQNFVVS-----GGSYLNKIKSGSTNTAFILKNNLTINAT 374

```


Db 800 T-----KQVDSLAQ-DGSLAGQISAAANTINTTGLITVEGNIKATISGLT----- 845
 QY 955 VDAGALRYKLVKNDGEFRLLHNPKEGELHNDLVRAEQERLEKQVEPACTQYGEKRY 1014
 Db 846 -----A-YNKAKLIDG-----TASGNRTVEVMAINAGSGSV 876
 QY 1015 RSRRAAARPPDTLPDQSLINALEKQELTAETOKSAKTKYRSKR--AVESDPLDDQ 1072
 Db 877 TAKTSSNV---NITDGLSTINGL-----NISNGRN---TVRLRGEIDVXYIOP----- 921
 QY 1073 SLFLEAALVETAPQOSEKCR-LAODEAKQKQKDLISRY--SNSAL-----SELAT 1124
 Db 922 ---GVASVEVIEAKKVLKXKQDLSDEEBETLAKLGVAVFEPNNALITVTOQEFTR 978
 QY 1125 VSNMLSYODELRLFDQAQSAWNTIACD 1154
 Db 979 PSSOYII-SEKACFTSSGNGAAYCTNAVD 1007

RESULT 6

US-09-806-709-37
 ; Sequence 37, Application US/09936709
 ; GENERAL INFORMATION:
 ; APPLICANT: LOOSMORE, Sheena M.
 ; APPLICANT: YANG, Yan-Ping
 ; APPLICANT: KLEIN, Michel H.
 ; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
 ; FILE REFERENCE: 1038-1138 MIS
 ; CURRENT APPLICATION NUMBER: US/09/806,709
 ; CURRENT FILING DATE: 2001-04-04
 ; PRIOR APPLICATION NUMBER: PCT/CA99/00938
 ; PRIOR FILING DATE: 1999-10-07
 ; PRIOR APPLICATION NUMBER: 09/167,568
 ; PRIOR FILING DATE: 1998-10-07
 ; PRIOR APPLICATION NUMBER: 09/206,942
 ; PRIOR FILING DATE: 1999-12-08
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 37
 ; LENGTH: 1222
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; US-09-806-709-37

Query Match 3.9%; Score 283.5; DB 5; Length 1222;
 Best local similarity 21.4%; Pred. No. 26-07;

Matches 268; Conservative 194; Mismatches 492; Indels 297; Gaps 65;

QY 40 PAENKG--KRTVGAQ-NIKVYKQGLVGTSMTRKAPMIDPSVVSRYVAALVENQYVS- 95
 Db 135 FEDKPGLSMLTTAGTIAVNNKKG-----FRDNTVITANGCGGSLKTYITFG 182
 QY 96 -----VAH-----NV-GYTDVDFGAGSNPDQHRFTYKLVKNNYKNDLHPEYEDYHN 143
 Db 183 NRDSNEETFRGRLLINISGKVDILMCAQCEMNRHWH-----GRSHWNTVLINVSNSYN 237
 QY 144 PRLKFEVTE---AAPIDMSNNNGSTSDPTKPKPERVYIGSRQPMKNDQKDDQYAA 199
 Db 238 VTTDSGSSASSPGAPLMAQSGNLGISSEKNDVF--NIAASSVNF-NIKPIVDKRTNG 294
 QY 200 YHVIAGNTHNRGAGNGSYLGGDYVRKAGEYGLPIAGSKDPSPMFIYDAEKQKMLI 259
 Db 295 NHLEFGN-----ISVLDG-----GMSFTFINASSSYOT 324
 QY 260 NGILREGNPFEGKENGQOLYRSYFDELHTS-LYTAANGVYITSGNDMGQSTIQ 319
 Db 325 YGVIIISQNFSA-SGSGSLKFKS-----EGSTYAA-FITKMDLILNATG---GNISLNG 373
 QY 320 KSGIPEIKITL-AMNSLPLKEDKXVNPBYDGNI-----YSPLNNGETLVFMQKXG 373
 Db 374 VAGIDSNLKKSLIAK-----KATTEGGNITLAADKKPTEIKGN---ITVKEG 418

QY 374 SLIFASDING--AGLIYEGNFTVSPNSNOTWOGACHIVSNGTV-----TW 419
 Db 419 ANVLTASANYGNDXSAISIRGN--VTNRGNLTVGSAINIEKMLIVESGAKFLANPNYSF 476
 QY 420 KTVGVEHDR-----LSKIG-----KGLTHVQKGN-----KQSLVGDGVILE 459
 Db 477 NYSGLFDMQGSNINISAKGALFKQIENTGSLNITTKDSNHRITIKNTINRKGDLNT 536
 QY 460 QQADDQKQKQAFSEI---GLVSGR-GTVQLNDKQFDTKRYF--GFRGGLDLN---GH 510
 Db 537 NNGDM-----TEIQGNISQKEGNLITSDKVNITRIITIKAGVNGDNSNPAISA 589
 QY 511 SLTF--KAIQVTDGAMIVNHTQAANVITIGNESIVLPGANNIKLDYRKIAVNGF 568
 Db 590 NLITIKELKLTND---LINSFKEKITAKNSMLITGDKSDAGNDAKVTFENVD 645
 QY 569 GETDKNKNRGLNLYKPTTEDRTLLSGGNLKDITQTKKILFFSRPPIPHVHLAK 628
 Db 646 SKISASDHRVTLNKKVZ-TSGDSTEDGANNNTG-LITIAKRVTVNNNTSH--KTVNI 701
 QY 629 RNSMEGIPOGEIYWDHMINET-----FRAENFOIKGSAVVSHPVSSLEGWNT--VSN 681
 Db 702 TASENVTTKAGTT-----INATGSVEVIAKAGDKIGSIEGSGGNVITASGDTIANVN 755
 QY 682 NMAITPVVPOQNTICTSDWTGTLTCQKDLITDKVINSIPKTOGNSINLTDNATN 741
 Db 756 ITGQVNTVAASAGAVTTTGSTINATGNANITTKGTEINGVVS--ASGVNIT--ASGN 812
 QY 742 VGLAKLNG-NVTLNHSOFTLSNNAIOGNIRLSPNST--ATVDNANL----- 787
 Db 813 TLNVSNIIGQNVYITANS-----CAITTEGSTINATIGDANITTOGNINGKV 861
 QY 788 ---NQNVELTDSAFSLKNSHPSHQ---IQGDKG--TVVLENAWTMPSPDTILO---N 835
 Db 862 ESSSGSVTLIATQ--TLAVNISGDTVITADKGLITGTSKINGKSVTSSGSGDIS 920
 QY 836 LILNNSITLINSAYKSSNNTPRRSLFETTTGSHRENTLVNKGKLSGQGTQFTSS 895
 Db 921 GTISNTVYSAGSLITQASKIEAKTGBANVTSA-----TGIGGTISG--NTVNVNAN 974
 QY 896 LFGYKSDKLYKSDAEGDIILSVNTSKE---PETLEQTLVESKQDQPSDKLEPTLEN 952
 Db 975 -----TDNLITIKQAR-----IKATGCAVTLTANGLITITETSSDITSSNGQTLITAKD 1023
 QY 953 DHVDAKALRYKLVKNDGEFRLLHNPKEGELHNDLVRAEQERLEKQVEPACTQYGEKRY 1002
 Db 1024 SSI-AGSI-----NANVTINITGTLITVAGSKIEAASGTLVINAKDAQLOGAASGDHT 1076
 QY 1003 -PLACVYGEPRKRSRARAARPPDTLPDQSLINALEKQELTAETOKSAKTKYRSK 1061
 Db 1077 VVNAITNANGSGSVIATTSRNVNITGDLITTINGLNTI--SKNKNVTLKGVIEDVXYIOP- 1134
 QY 1062 RAVPDDPLDQSLFALAALEVIDAPQOSEKCR-LAODEAKQKQKDLISRY--SNSAL 1118
 Db 1135 -----GVASVNEVIEAKKVLKXKQDLSDEEBETLAKLGVAVFEPNNALITVTOQEFTR 1180
 QY 1119 SELATVNSMLSYODE-----LDLFDVDAQ-----SAWNTIADQKR 1156
 Db 1181 -----TINT-----QNEFTTRPLOSQVITISEKGVCFILNGAVITCTINADIAR 1222

RESULT 7

US-09-806-709-34
 ; Sequence 34, Application US/09806709
 ; GENERAL INFORMATION:
 ; APPLICANT: LOOSMORE, Sheena M.
 ; APPLICANT: YANG, Yan-Ping
 ; APPLICANT: KLEIN, Michel H.
 ; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
 ; FILE REFERENCE: 1038-1138 MIS
 ; CURRENT APPLICATION NUMBER: US/09/806,709
 ; CURRENT FILING DATE: 2001-04-04

```

; PRIOR APPLICATION NUMBER: PCT/CA99/00938
; PRIOR FILING DATE: 1999-10-07
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 09/206,942
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-806-709-34

```

Query Match 3.9%; Score 283.5; DB 5; Length 1228;

Best Local Similarity 21.4%; Pred. No. 2e-07;

Matches 268; Conservative 194; Mismatches 492; Indels 297; Gaps 65;

```

QY 40 EAENKG-KFTVGAQ-NIKVYKQSQQLVSTSTKAPVDPVSYSNGVALVYNQYTVS- 95
DB 141 FEDKPELSNLTITAKGTIAVNNKG-----FREDVTLNKGQGLSEKLTITG 188
QY 96 -----VAH-----NV-GYTDVDFGAGNNPDQREFTYKIVKNNYKKDLHPYEDDYN 143
DB 189 NRDSNFEHFGRGLNLSGKVDILMQAQENNRHW-----GRSHMVTILNVSNSYFN 243
QY 144 PRLHFTVE-----AAPIDTSMNNSSTSDPKYPERVIGSGRFNRNDQDGQVAGA 199
DB 244 VITDSSSASSPGAGLNMOSGLNLSNNDTVF--NIASSAVNF-NIKRPVLDVTVNG 300
QY 200 YHLYPAKTNHORGAGNGSYVLGQVRKAGEYGP-LPIAGSKDSGSPFIYDAEKQWLI 259
DB 301 NHTLPKKN-----ISVLGG-----GMSTFFNASSSNVOT 330
QY 260 NGHLESPPEKCKENGFQVRSYSYDEIFERDLHLSLYTRAGNGVTTISGNDGQSITQ 319
DB 331 YGVILISQNFSA-SGSSSLKFS-----EGSTHAA-FTIKKIDILNATG--GNLSNQ 379
QY 320 KSGIPSEIKITL-ANMSLPLKKKQKQVNPVYDGNP-----YSPLAKSGTILYFDQKQ 373
DB 383 VAGIDSEIKKSLIAN-----KNITFEAGNTTLAADKKPEIKKN---ITVDEG 424
QY 374 SLIFASDINOG--AGGLYEGNFTVSPNSQVQAGLHVSENVTV-----TW 419
DB 425 AAVTLRSANYGDKSLISRG--VTMKGLVTVGSALNIEKRLTVGSGAKFLAANPSF 482
QY 420 KNGVEHDR-----LSKIG-----KQTLHVQAKGEN-----KGSIVGDKVILE 459
DB 483 NVSGLEPNQOKSNISIAKGAIFQDIENGTSLNITKSDSNHTTIKKNITNRKGLDMLIT 542
QY 460 QOADOQKQAFSEI---GLVSGR-GTVQJLNDKQDDDKYF--GPRGRDLN---GH 510
DB 543 NNGDN-----TEIQGNISQKEGMLTSSDKNITERTTTIAGVNGDSNSBATS 595
QY 511 SLTF--KRIQNTDEGAMVNNNTQAANVTITGNSIYLPNGNNINLDPREKILANCMF 568
DB 596 MLTIKTELKLTND-----LNTSGFNKAETIAXONSMLTIGNSDAGNTDAKVTPSNKD 651
QY 569 GETDKKNGRLNLYKPTEDRLLISGNTAKGDIQYQKLFPSGAPTPHAYNLNK 628
DB 652 SKTASHENVTLSKYE--TSGDITDSTEDGNNNTG-LITLAKVTVNNITSH--KTVNI 707
QY 629 RWSMEGIPQGEIWDHMINRT-----FRAENFOIKGASVSRVSSISEGNTV--VSN 681
DB 708 TASENVYTKAGT-----INATGSVEYTAKGDIKGLISNSGNNTTASGDLVNSN 761
QY 682 NANAATFGVNPQNTCTIRSDWTGLTTCQVVDLDTKVINSIPTQINGISILNATNATN 741
DB 762 ITGQNTVTAASGAVTTTKGTINATTGNAANTTTGSEINSEVYS--ASGNVIT--ASGN 818
QY 742 VKGIKLNG--NTLNNHSGFTLSNATQIGNTRLSDNT--ACTDNANT----- 787
DB 819 TLNVSNTIGQNTVTVANS-----GALTTEGSTINATTGANTTTQTNINGKV 867

```

```

QY 788 ---NGNVHLTDSAQFELKNSHFSHQ---IQDKG---TIVLENNATWMPSDTLTQ---N 835
DB 868 SSSSSSVTLIATGQ-TLAVGNISGDTYITADGKULTQTSKINGKSTSSQSGDLS 926
QY 836 LILNNSITILNLSAVSASSNNTPRRRSLETTPTTSAHERFNTLVNKLSGQGTQFTSS 895
DB 927 GTISGNTVSVASGSLTTOAGSKIEAKTGEANTSA-----TGIGGTLISG--NTVAVN 980
QY 896 LFGYSDKGLKISNDAGDYILSVNTKE---PETLEQLTVESKQNDPLSDTKRTYLEN 952
DB 981 -----TNLTIYKQAR-----IKATGAVTLTATGGLITETSSDITSSNGQITVLAKD 1029
QY 953 DPAVDAALRYKLVNDGFERLHNTKEQELHNDLVAAEQERTEAKQVE----- 1002
DB 1030 SSI-AGSI-----NAAVNTLNTGLITVYASKTEAASGLVIMADQADGAASGHT 1082
QY 1003 -PTAKTQTEPRKRSRARAAPDTLPDQSLNALMAKQELTAETQSKAKTVKVSX 1061
DB 1083 VNAITNANSGSVTATISSRVNITGDLITINGLNTI-SKNGKNTVLILKGVIEDVRYIQP- 1140
QY 1062 RAVFSDPLDQSLFALEALAVIDAPQOSEKDR-LAQBEEKQKQKDLISRYS--NSAL 1118
DB 1141 -----GLASVNEVIAKRALRYKQVLSDEKETTLAKGVASVAPALPNNAI 1186
QY 1119 SELSATVNSMLSVODE-----LDELTVDAQ-----SAWNTNIAQDKR 1156
DB 1187 -----TLNT-----QNEFTTRPLSGVTTSEKVCFLINGATITCINIADIER 1228

```

RESULT 8

US-09-806-709-67

Sequence 67, Application US/09806709

GENERAL INFORMATION:

APPLICANT: LOOSMORE, Sheena M.

APPLICANT: YANG, Yan-Ping

APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH

TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS

FILE REFERENCE: 1038-1138 MIS

CURRENT APPLICATION NUMBER: US/09/806,709

CURRENT FILING DATE: 2001-04-04

PRIOR APPLICATION NUMBER: PCT/CA99/00938

PRIOR FILING DATE: 1999-10-07

PRIOR APPLICATION NUMBER: 09/167,568

PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: 09/206,942

PRIOR FILING DATE: 1999-12-08

NUMBER OF SEQ ID NOS: 95

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 67

LENGTH: 1536

TYPE: PRT

ORGANISM: Haemophilus influenzae

US-09-806-709-67

Query Match

Best Local Similarity 3.9%; Score 283.5; DB 5; Length 1536;

Matches 244; Conservative 162; Mismatches 400; Indels 357; Gaps 56;

```

QY 42 ENKGRFTVGAQINIVNKKQQLVGSMTKRAMIDFSVNSKNGVALVYNQYTVSVAHNG 101
DB 219 KNGGVISVNGSISL--AGQKITISDIINFTIYTS-----AAPNE-----AVNIG 264
QY 102 YTVDFGAEAGNPPQHRFTYKIVKNNYKKDLHPYEDDYNPRLKFTVTAAPIDPTSN 161
DB 265 -----DIKAKGINVRAATI-----RNO-----GKLSADSVSKQSGN 298
QY 162 KNGSTVSDPTKYPERRVIGSGQFWRNDQDGQ--VAGAYHLYPAKTNHORGAGNGYS 219
DB 299 IVLSAKEGHAE-----IG-GVISAQNGQAKGKMLITGDKVTLTKTGAVIDLSGKEGGET 351
QY 220 YLGQVRKAGEYGPLPIAGSKDSSGSPFIYDAEKQ---KA---LING--ILREKNPFE 270

```

Db 352 YLGGDERGEGKGIOLAKKTSLKSGSTINWSGKGGRAIYWGIALIDENIMAGSSDI 411
 QY 271 GGENGFOLVRSYDELFERD-----LHIS--LYTRAGN 302
 Db 412 AKTGSGF--VETSGHD--LFIKONAIYDAKEMLDPONVSIINAGRSINSEDEYTSGN 468
 QY 303 GYVITISGNDNGGSIYOKSGIPSEIKITILANMSP--LKECXVH--NPRVDGNIT-- 355
 Db 469 SASPYKXKE-----KTLINLTLESILKGTFFNITAMOR-----IYVN 508
 QY 356 -SPRLNGEITLYPMOKOSLIEMSDINOGAGLYFEGNFTVSPNSQTMOGAGIHVSEN 414
 Db 509 SSILNWSGLTIWSEBSG--GVEIENDITGDDT--RANLITVSG 552
 QY 415 STYTWKVVGVHDLRSKIGKTHVOKAGSKNKS:SYGDGVYIIIEQADDQNKQAFSEI 474
 Db 553 GWVD-----VHGX:SLGQGNINITAK-----QDIAFKGSNQ----- 585
 QY 475 GLVSGRGTVOLNDKQFDTEKVFYFGGRLDNGH--SLTEPK--IQNDEGAM 525
 Db 586 -VITSGGFTTSGNQK-----GFRPNVSLNKGSGLOFTTKIKTKALTNKPESTL 635
 QY 526 IYVHNTIQAANVTITGNSIYLPNGNINKLDYRKELIAYNGMFEETDKN-----XANGR 579
 Db 636 -----NISGXNISMVLP-----KHESGYDKFKGRTYMWLTSLNVSSEGE 675
 QY 580 INLIYKPTEDRITLISGNTLKGDIQTKGLFFSGGFTPHVNHANRMSBEGIPQG 639
 Db 676 ENL-----TIDSR-----GDSAGTLTO-----PYN-----ANGISFN 703
 QY 640 EIVWDHDMINRFEKAB-----NFOIK--GSAVAVSRVSSIEGNTVYSNANANPFPVP 691
 Db 704 K-----DITFENYERANRVPDIAP:PIGKIKYSIANYASFPNGNISVSUGGSVDFTLLA 755
 QY 632 NQOQNT--ICRSMTGLTTCQV-----DCTDYKINSIPTQUNGS 731
 Db 756 SSSNVQTPGVVINSKYFVSVSSLSRFTSGSTKISIEKDLTINATGNTLLQYBGT 815
 QY 732 INLTDN--ATANV-----KGLAKINGVTLTNHSGFTLSRNATQION-- 771
 Db 816 DGMIGKGIYAKKNIIFBEGNITFGSKRAVTEIEGVITINNA-----NVTLLISDEDMH 869
 QY 772 -----ILSGNSTATVVDNANLNGVH:LTSAOFSILKSHFSHOIQ--DKG-TT 817
 Db 870 QKPLT:KKDVI--NSGNLILAGNIVNLAGNLVRESNANFK-AI:NTFNVVGLPUNKGNSN 928
 QY 818 VTLLENATWMTMPSDTTLQMLTL-NRSTITLNSAYASSNTPRRSRLETTETTSAR-- 872
 Db 929 ISIAKGAARFKOINDSKNLSITTSSESTYRTIISGNITNKGDLNITVEGSDTPEMIGGD 988
 QY 873 -HREFTLYNG-KLSGGGTFFPSTSLPEVYS-----KLIKLSND----- 909
 Db 989 VSGKEGNLTISDKINTTKOITTKAGVUGENSDATNNAKLTITKELKLTODLINSGF 1048
 QY 910 -----AESDYILSVNRTGKEPET-LEQTLVESKQNOPUSKLFETLENDDHDAALAY 962
 Db 1049 NKAELTAKGSDLTIGTNTSADGTNAKVTYNYQKOSKISADGHVTLHSVETISSNN 1108
 QY 963 KLVKNDGEFRLHPIKEQELHNDIVRAEQARTIEAKQVEPFAKTQGEPEVRSRPARA 1022
 Db 1109 TEHSDNNAGLITDAKVVVYNNNITSHKAV--SISATSEIITTKGT-----TIVALTG 1160
 QY 1023 APFDTLPDOSLNLNLEAKQAEILT 1045
 Db 1161 NVEITAGTSGIIGGIESSGSVT 1181

RESULT 9
US-09-806-709-71

Sequence 71, Application US/09806709
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.

APPLICANT: YANG, Yan-Ping
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
 FILE REFERENCE: 1038-1138 MIS
 CURRENT APPLICATION NUMBER: US/09/806, 709
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: PCT/CA99/00938
 PRIOR FILING DATE: 1999-10-07
 PRIOR APPLICATION NUMBER: 09/167,568
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 09/206,942
 PRIOR FILING DATE: 1999-12-08
 NUMBER OF SEQ ID NOS: 95
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 71
 LENGTH: 1477
 TYPE: ERT
 ORGANISM: Haemophilus influenzae
 US-09-806-709-71

Query Match 3.8%; Score 280; DB 5; Length 1477;
 Best Local Similarity 20.2%; Pred. No. 4e-07;
 Matches 230; Conservative 151; Mismatches 402; Indels 358; Gaps 49;

QY 42 ENKGFYVGAONIKYVYKQOLWTSMTKAPMIDFVSVRNGVAALVENQYIVSAHNG 101
 Db 219 KKEGVISVNGGISIL--AQKILISDILNPTIYGI-----AAPENE--AVNIS 264
 QY 102 YVDVPGAEKNNPQHRFTYKIVANNYKDONLHPYEDDYNHRLHFVEAAPTDMTSN 161
 Db 265 ---DIFAKGNINVRATIT--RMQ-----GKLSADSVSKXSGN 298
 QY 162 MNGSTYSRTKYPFRVIRIGSGQFWRNDQKGD--VAGAYHYL:LAGNTNQGAGNGS 213
 Db 299 IYLSAKGEAL-----IG-GTISQVQAKGAKM:ITGDKTTLTGAVIDLSGKEGPT 351
 QY 220 YIGGVYKAGBYGRLPIAGSKDGSMPFIYDAKO--KW-----LINS-ILREGNPE 270
 Db 352 YLGGDERGEGKGIOLAKKTSLKSGSTINWSGKGGRAIYWGIALIDENIMAGSSDI 411
 QY 271 GGENGFOLVRSYDELFERDLHISLYTRAGNYTISND-----NQG 315
 Db 412 AKTGSGFVETSGHLYLSDNALVTKKEMLDPDV--TIAEDPFLNNTGINDPFTGTGEA 470
 QY 316 SITOKSGIPSEIKITILANWSLPIKEDKVANPRYDGPVYSPRLNGEITLYMDOKQSL 375
 Db 471 SDPKKN--SEKTTILNNTI-----SNYKXNMTNITASRK--L 506
 QY 376 IYASDINQA-----GALYFEGNFTVSPNSQTMOGAGIHVSENSVTYTRVN 422
 Db 507 TVNSINIGNSHLILHSKQRGQGVQIDEDITSKGNLTIYSGWVDVHKUL----- 560
 QY 423 GVHEDRLSKIGKTHVQAKGENKGSISYGDQVILBQADDQNKQAFSEI:GLVSGRT 482
 Db 561 -----LDQOFNLITTA--ASVABEG--NNKADANAK--TVAQGT 595
 QY 483 VOLNDKQFDLTKYFGRGGRLDINCHSLTFKRLQNTDEGAMIVN-----HNTTQAN 536
 Db 596 VITIGBCK-----DEFRANVSLNG-----TGKGLNISVANNLHNLSGIN 637
 QY 537 VT--ITGN-----SIVLPNGNINKLYRKE----- 561
 Db 638 ISGNITLNTQTRKNTSYQTSYDHSNMVSAALMIFGAMFTFIKYSISNKGILTYGSSA 697
 QY 562 -IAYGWFEETDKN--KHNGRLNLYKPTEDRT-----LLSGGNTMKGDITOT 608
 Db 638 GVNPNNGNMSFNLKBAKYNFKLEKNNMTS:KPLFIRPLANTATG----- 747
 QY 609 KGKILF-----SGPTEPHAYNHLN-----KRWSEMEGIPQGEIYWD----- 644
 Db 748 -GSVFEDIYANHSRGALMKSEINISNGAFILNSVGRDDAFKINDLTINATNNSNFS 806

QY 645 -----HMINRTPKAEHPQIKGSAVY--SRVSSIEGNTVSNNAK-TEGV 689
 Db 807 LQOTKDFYDGYARNAINSTY--NISIIIGAVTLIGSSSSSITITNIIIEKAANTLLEA 863
 QY 690 --VINOONITCTRSWTLGTLTQKVDLTDTKVINISIPRTQINGSI--NLTNCAVANKGI 745
 Db 864 NNANNOON-----IADRYIKGSLVNGSLSTIGENDIKNLITSESAITPKGK 912
 QY 746 AK-----LNGAVTLTNHSOFTLSNNAQIGNI-----RLSDNISTATVDMANLNG 789
 Db 913 TROTLNITGNFTNNGTAELINITOGVVKLGAVTNMGDLNITTHAKRHOUSIIIGDIIINKG 972
 QY 790 NWHLTSAQFSLKNSHSPHOIGDKGTVTLENATWTPSD--TTLOKLTNNGSTILNLS 847
 Db 973 SLMTIDS-----NNDADIIQIGN-----ISOKEGNLTISDK.NITKOITIKKIDGESS 1022
 QY 848 AYGASSNTPRRASLEFETTPPSAHERFVTLTVNGKLSGGTGFQPTSSIFPKSDKLKLS 907
 Db 1023 SSQATSNANLTIKTKELTLEDLSIGFVKAITAK-----DGRDLITIG 1066
 QY 908 NDAEGDYI.SVANTGKEPETLEQLTVESK-----DNQPLSDKLPTELDHVDAGALRY 962
 Db 1067 NSWDG-----NEGAETVTVTNKOSKLSADGRVTLNISKVTSSSNGSRS----- 1114
 QY 963 KLVKNDGEFLNHPKQELHNDLVRAE-----QAEPTLEAKQVEPTAKTONG 1010
 Db 1115 ---NSDNDTGLITAKVVEYVKODISLKTNTVITASEKVTTTASGTITNATKASITTKIG 1171
 QY 1011 E 1011
 Db 1172 D 1172

RESULT 10

US-09-806-709-53

/ Sequence 53, Application US/09806709
 / GENERAL INFORMATION:
 / APPLICANT: LOOSMORE, Sheena M.
 / APPLICANT: YANG, Yan-ping
 / APPLICANT: KLEIN, Michel H
 / TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 / FILE REFERENCE: 1038-1138 MIS
 / CURRENT APPLICATION NUMBER: US/09/806,709
 / PRIOR FILING DATE: 2001-04-04
 / PRIOR APPLICATION NUMBER: PCT/CA99/00938
 / PRIOR FILING DATE: 1999-10-07
 / PRIOR APPLICATION NUMBER: 09/167,568
 / PRIOR FILING DATE: 1998-10-07
 / PRIOR APPLICATION NUMBER: 09/206,942
 / NUMBER OF SEQ ID NOS: 95
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 53
 / LENGTH: 977
 / TYPE: PRT
 / ORGANISM: Haemophilus influenzae
 US-09-806-709-53

Query Match

Best Local Similarity 3.8%; Score 277; DB 5; Length 977;

Matches 241; Conservative 164; Mismatches 436; Indels 308; Gaps 52;

QY 123 IYKRVNYKONLHPYEDDYHNPRLHKFVTLAAIDMTSNMGSYSRTYPERVRISSG 182
 Db 17 VYRGSTIEKKA-PTTITTHASISIEGSLMGAVNISANKVWTLTINVYGAALTISE 75
 QY 183 RQFWKNDQKGDQVAVAGVHTJAGN-----THNORAGNGY-SYIAGEVAKAGEY 231
 Db 76 RD-----GVEINGNITSEKGNLITKAGSVVYHKNITLIGSPINTISGD----- 120
 QY 232 GELPIASGKDGSSPEFTIYDAKOKMLINCIILRENGPEEGKNGFQIVKSYFDEIFERD 291

Db 121 ----IAPEKGNLIT-----ITAQGNITSNKD-GKOL-----EL 148
 QY 292 LHTSLY-TRAG-----NGYITISNDNGQSSIIQKSGISEIKITLTMGLPKEXD 342
 Db 149 NNVSLNIGTAGLAFIANONNFTNINISGALINSVLT-----INQITKKA 193
 QY 343 KYHNPRIYDGP-NIYSPRLNNGSETLYFMDQKGLIIPASDINOAGLYEFGNFTVSPNSN 401
 Db 194 KAMNTSYDSYWNVSTLITLSDAKFTFIK-----YVDSNISTNSDS 234
 QY 402 QTWQAGIHSNENSTYWKV--NGVEHRLSKIGKGTILVQKENGNSISV-GDKVTL 458
 Db 235 RSFAGVKEH-GKONKEKFWIGNNAKAEFLPKPEKTTIPRPILOPTLSNISTVGGSVFE 293
 QY 459 EGOAD--DOGNKQAFSEIGLYSGR-----GTQVANDDQPD-----TD 494
 Db 294 DLYANLMKGTLEIKMDSINVSNGSLTINSHVARKYNAFEINNDLTITATNSNFKLQSTD 353
 QY 495 KFYEGFR-----GGRDLNGH-----SLTEKION-----TDEGAMIV 527
 Db 354 SFERNGYRNALINSTANISILGONVTLGGQSSSSIMENIIIEKAANTLEADNSHSDV 413
 QY 528 NNTTQAANTTTGNSIYLPKANNINKLDYKELAVNGWAGETDKRKNAGRLNIIKYET 587
 Db 414 KORTINLGNLTVEGNSLIGERNANINGLSIEKEAIFG--KTKDSLITGNFTNGT 469
 QY 588 TEDRTLLSGGTILKGDITOTKGLFPSGRPHVYHNLKRWSEMGIPQGGIYWDWD 647
 Db 470 AE---INISQVSLGDIIT-NDGL-----NITTHAK-----SGQSIIRGDIINKGN 514
 QY 648 INRTPKAENPQIKGSAVSNVSSIEGNVTVSN-ANATGV-----VPOONTICTRS 701
 Db 515 LNTIDNNSNALEIIGG---NISQKGNLTISDKNTVITQILITAKGVDESSSSSIS 569
 QY 702 DWTGTLTQCKVDLTDTKVINIS-PRTOJNGSIN-----LITNATH-----NYGLAK 747
 Db 576 DANLITITKELTITLNLINISEFNPAETAKNDSLLIGKASDNSNAKQVTEKDKOSKI 629
 QY 748 LNGAVTLTNHSOFTLSNNAQIGNIRLSDNSTATVDNANLNGVHLTDSAGSLKNSHF- 806
 Db 630 SAGHNVTILNKKVTSN-----SDSGTG---NSDNNIGLITISKQVTVNSIT 676
 QY 807 SHQ-----IQDGKTTVTLENATWTPSDTLTONTLTNNGSTITLNSAYSNASNT 856
 Db 677 SHKVTNISASBGGITTKAGITIVATTGSEVETAKGDISGISTIVS---TASIGDL 732
 QY 857 PRSSLETTPTPSAHERNTLTVNGK--LSGGTGFQPTSSIFPKSDKLKLSNDAEGDYI 915
 Db 733 TVRKALATISATGAA-----TLTATGNTLTTEAGSSITST-----KGVDLSAQ-DGSLA 781
 QY 916 LSVNTEGKEPETLEQLTVESKNOQLSDKLFTELDHVDAGALRYKLVKNDGEFLN 975
 Db 782 GOISAAVNTLNTTGLTLTVEGSLNIKATSGTL-----AIYADAKLID- 823
 QY 976 PIKQELHNDLVRAHQARTLEAKQVEPTAKTQVEBKVRSPRAAPAPPTLPDQSLN 1035
 Db 824 -----TASGRTEVNATWASGGSVYAKTSSNV---NITGDLSTIN 861
 QY 1036 ALEKQAELEAETOKSAKTKVYRSK--AVRSPELDQSLTELEAALEVDAQDSEKD 1093
 Db 862 GL-----NISMNGN--TVRLKGEIDIVKIQD-----GVASVEEVLKARVLEVY 906
 QY 1094 R-LAQBPAKROKROKDLISRY--SNSAL-----SELSATVNSMLSVODELRLFPVDAQS 1145
 Db 907 KDLSDERETLAKGVSAVFEPPNALITINQNEFTTRSSQVII--SEGAACSSNGKA 965
 QY 1146 AVYNTINAD 1154
 Db 966 AVCTNAD 974

RESULT 11

US-09-806-709-51

```

; Sequence 51, Application US/09806709
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Steena M.
; APPLICANT: YANG, Yan-ping
; APPLICANT: KLEIN, Machei H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1138 MIS
; CURRENT APPLICATION NUMBER: US/09/906,709
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: PCT/CA93/00938
; PRIOR FILING DATE: 1999-10-07
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 09/206,942
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 51
; LENGTH: 983
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
; US-09-806-709-51

```

```

Query Match      3.8%; Score 277; DB 5; Length 983;
Best Local Similarity 21.0%; Pred. No. 3.4e-07;
Matches 241; Conservative 164; Mismatches 436; Indels 308; Gaps 52;

```

```

QY 123 TVKNNYKDNLPYEDDYHNPRLKHTVBAFIMTNNMGSTYDRKTYPERVRIQSG 182
DB 23 VVRSSTLEKNA-PKTTTHAGSTIEOSIKKGAANYISATKNVNTDIVVNGALTLHSE 81
QY 183 RQFWRNDQKQCVAGAYHYLTAGN-----THNQGAGNGV-STYGGDPYAKKGEY 221
DB 82 RD-----GVINGNITSEKGNLTIRAGSWVYHKNITLSEGLNTTSGD----- 126
QY 222 GPLLIAGSKGSGSPMFLYDAEKQKWLINGILREGNPEKGFOLYVEXYFDEFERD 291
DB 127 ---IAPEKGNMLT-----IAQGNLITSNKD-GKQL-----RL 154
QY 222 LHTSLY-TRAG-----NGVTTISGNNGGSGITOKSGIPSEIKITLANKSLPLEKQ 342
DB 155 NNVSLNCGAGLNFANONNPTNHSICATINISGVV-----INQTTKNA 199
QY 343 KYNFRYDGP-NIYSPRLNGETLYFMDQKQSLIFASDINGAGGLYEPGFTVSPNSK 401
DB 200 KAWMTSYDSYVAVVSTLTLSDNAKTPIK-----YDSNHSCTMSOS 240
QY 402 CTWQAGIHSSENSVTWKV--NGVEHRLSKIGKGLHVOAKENKGSISY-GDGKVTIL 458
DB 241 RSFAGYFTH-GKCNEMKFNIGNNAKAFRLKNEKTPNRPFLPIQLANISVYGGGSVFP 299
QY 459 EQQAD--DOGNQASSELGYSGR-----GVQLANDKQPD-----TD 494
DB 300 DIYANLKGKTELRKMSDINVSQSNLTLSHYKXNAFEINKDLTINTNSFNLRQTSJ 359
QY 495 KEYFGR-----GGRLDINGH-----SLTEKIQV-----TDEGAMTY 527
DB 360 SFRRGRNNAINSTHNISLIGAVVTLGGNSSSSIMKNIITIKAAVTELEADNHSNDV 419
QY 528 NHATTCAAVTTTNGESVLPNGANNIKLDYRKELIAYNGFEGEDTKRNGRLMLYKPL 587
DB 420 KDRITMLNLTVEGNLSLIGENANINGNISLEKKA-FKG--KTKDSLNTITGFTNNGT 475
QY 588 TEDRTLLSGTNLKGDLTQTKGLFFSGRPTPHAYNHLMKRSEMEGIPGELVWHDHW 647
DB 476 AE---INISQVSVLADIT-NDGKL---NITTHAK-----SQKSLIIGDITNKQCN 523
QY 648 INTFEAFENFOIKGSAVSRVSVSIEGNWTVSN--ANATFGV-----VPMOQNTICTRS 701
DB 521 LNTDNNNAHEIGG-----NISCKEKNLTSSDKXNITKQITIKAYVDESSSSSTAS 575
QY 702 DWGTLTCKKVDLTDITVINSIKPQVINGSIN-----LTDNATA-----NVKGLAK 747

```

```

DB 576 DAULTTKTELTLDNINISGNKKEITAKNDSLIIIGKASSNSNSAPQVTFDYKXSKI 635
QY 748 LKNGVTLTNHSQPTLSNNAQOINIRLSDNSATVDNANNGVHLTSAQFSLKNSHF- 806
DB 636 SAGNHVTLANSVETSN-----SDGSGT---NGSDNNIGLITSAQVTVNSNLT 682
QY 807 SHQ-----IQNGKTTVLENAATWMPSDTTLQNLTLNNSITTLNAGSASNNMT 856
DB 683 SHRTVNISASEGTLTLAGTTINATGASVETAKTGDISITSGKTVS-----TASIGDL 738
QY 857 PRBSSELTETTPSAHERFVTLVWCK-LSQGTFOPTSSLPKYSCLKNSABEDYI 915
DB 739 TVKKAATISATEGA---TLATGNTLTTEAGSSITST---KQGVDSLAC-DOGSLA 787
QY 916 LSVANTGKEPTELRLQTLVESKNOPLSDKLTLENDYDAGALRYLVGNQDEPFLAN 975
DB 788 GQISAAVNTLMTGTTLTVGSSNIAKATSGTL-----ALNADAKLDO----- 829
QY 976 PIKEDELHNDLVRAEQERTLEAKQVEPTAKTQTEPKYRSRRARAAPDPTLPDQSLN 1035
DB 830 -----TASGNRTVNAATNAGSGSVTAKTSNV---NITGDLSTIN 867
QY 1036 ALFAKQBELTETOKSAKTKKYSER--AVPSDPLDQSLFALMALETVIDAPQOSEKD 1093
DB 868 GL-----NISENERN--TVRLKKEIDVRYIOP-----GVASVEVTEARVLEKY 912
QY 1094 R-LAQEAERKQKQKOLISRY--SNSAL-----SELSATVNSMLSVQDELDRLFPVQAS 1145
DB 913 KDLSEBETLAKLGVSAVRFEPNNAITINTQNEPTTPSSQVIL-SEGACACSSGNA 971
QY 1146 AAVTNIAOD 1154
DB 972 AVCTNVAD 980

```

```

RESULT 12
PCT-US04-07301-693
; Sequence 693, Application PC/TUS0407001
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF MONTPEABLE H. INFLUENZAE
; FILE REFERENCE: 2835/38815A
; CURRENT APPLICATION NUMBER: PCT/US04/07001
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 693
; LENGTH: 1381
; TYPE: PRF
; ORGANISM: H. influenzae
PCT-US04-07001-693

```

```

Query Match      3.8%; Score 275; DB 1; Length 1381;
Best Local Similarity 19.1%; Pred. No. 6.8e-07;
Matches 274; Conservative 192; Mismatches 484; Indels 484; Gaps 62;

```

```

QY 42 ENKGFVGAQNTIKYVNGQGLVGTSMTKAPM-IDFSVSRNGVAAVLENGYISVAHVNG 101
DB 108 KNEGVISVNGSISIL--AGQKITISDIINPTITYSI-----AAPEHEAT-----NLG 153
QY 102 YTDVDFASGNPDQGRFTYKIVRANNYKKNLHPYEDYHNPRLKHTVBAFIMTNS 161
DB 154 ---DIFAKGNIN-----VPAANIRNG-----KLSADSVSDKSGN 187
QY 162 MNGSTYSDRTKYPERVRIQSGRQFWRNDQKQDO--VAGAYHYLTAGNHNDRGAGNYS 219
DB 188 IYLSAKGEAF-----IG-GVIAQQAQKQKAMITGDVTLTKTGAVYIDSGKGGSET 240
QY 220 YIGDVRKAGEYGPPLIAGSKGSGSPMFLYDAEKQ-----KN-----LINGILREGNPPG 271

```

Db 241 YLGGDERGKNGIOLAKTISLEKSTINVSKEKGRATWGDIALIDGINAOGD-1A 300
 Qy 272 KENGFOUAKSFEDEFERDLHTSLYTRAGNGVTTISGNDGSGITQSGIPSEIKITL 331
 Db 301 KTGGF-----VETS-----GHYLSIGND--AAVEAKEMWLPDQVIT 335
 Qy 332 ANMSLEPKEDKVNPRYDGN-----IYSPILNGETLTFMOKOGSLIFA 378
 Db 336 SNGN---DDQSOLKDRDSDPNKILADNKHVYNNKLTSTALAKGIVNISAKKCVNTAD 392
 Qy 379 SDINOGA-----GGLYFEGNFTVSPNSNOTWQ--GAGIVSNSNV----- 417
 Db 393 INHNGTLTLHSEOGGYEINGJLTSBONAKLTIRAGSWVYVHAKNTITMGFLNITAGGSV 452
 Qy 418 -----TWKNGVEHRLSKIGKGLTHVQAKENGK 447
 Db 453 APERAGDGRASAPAKIYAQVITAGSGODFRNNVS---INGRGGLKFTAKG--NGG 508
 Qy 448 SISVG-----DGKVLIEGQAD-----QNKQAFSEIGLVSGRGTQ 484
 Db 509 NFSKEDGYLANSINISHTANNQSLYFHRQGYTWNLTOLNVDSDSFLTSIXDAIK 568
 Qy 485 IN-----DDQ-----FDTKFEFGGRLDNGSLTFKRIQ-----TD 521
 Db 563 VGGYDAKDKKNTGGIFTRDITFNVKQARVDIS--YTLPISPVKNRSLAANPDGNITV 627
 Qy 522 EGAMTV-----HNTTOANVTITG----- 541
 Db 628 KGGGVNKLKFNALSNNKPEGVNISRPINVTESQJLNTGSPSTLEWVANDLITNAT 587
 Qy 542 -----NESIVLPCNG-----NINKLDYKKEIAY---NOMFE 570
 Db 688 NSFVSIKEIGETCHLDITGLKYNANVTIKGSNVTLLSGNNAKTYEPKNTVEKANITLAS 747
 Qy 571 TDKNHNGSLNLYKPTEDRTLLSGTNLKGDIT-----QYKGLPFSGRTPH 621
 Db 748 ANFGHKGALITVAGNITQKLVATGDTIIVSDGFTVGDAPFNNTNMLNITGFTTN 807
 Qy 622 AYV-----EAKRSEWBGIPQGEIYMDHMINETFAENF 657
 Db 808 GSIIDVKKGAUKGNITNEGSLNITTHAN--TKQKITLITGNITKKGDLNIRNKNNA 864
 Qy 658 QIKGSAVYSHVSSIEGNTVVS-----NNANATPQVNOQNTICTR- 700
 Db 865 ELOIGG---K-SQKRGMLTSSDKVNTTKQITTAGVNGENSDSGTENNANMLITKXT 919
 Qy 701 -----SDWTGLTQOKV---DLTDFKV----- 719
 Db 920 LELTNLNLISGFHKAETIAKNSDILIGKASSDSGNAAGQVIEDKXKDSKISAGNHTV 979
 Qy 720 INSIRKOINGSINTDNATIANVGL-----AKLMGNVTLTNHSOFTLSNNATQIGIR 773
 Db 980 LNSEVEVS-NKSNNAAGDSNGNMGTLISAKDVAVANNIT--SHKTIINISAT--TGNVT 1033
 Qy 774 LSDNET-----ATVDNANINGVNLITDSQPSLKNSHPSHQIGDKGT--VYLEN 822
 Db 1034 TKEGTTINATGVEVETATGDIKIGIR-SKSGGYTLITAGDTLAVANISGNTSVTANS 1092
 Qy 823 ATWMPSDTTLQNLINNSTITLINSAYSASNNTPRRSLETETTPSSEHPNTLYTNG 882
 Db 1093 GTLITKADSTIKG--TGSVTLTISGSDIGGITSKTVSA--TALT-----DGLTYKG 1140
 Qy 883 --KLSG--OGTFEFTSLRPGYKSKDKLSNABEDYLSV--ANVGE-----PEYLEQ 930
 Db 1141 GAKINATEGATLITASSGKLTTEASSITSAKQVUDLSARBNGISGINAANVTANTGT 1200
 Qy 931 LTLVESKDNQPLSDKLEFLNHDVADAGALRYKL VANDGEFLNHPIKEGELHNDIVAB 990
 Db 1201 LTTVKGSSINANGTIVINADALDGT----- 1229
 Qy 991 QABRTLEAKOVEETAKTQTEPPKVRSRARAAAFDTLLPDOSLIMALE--AKQAEIJAET 1048
 Db 1230 SGRDVT---VNAVNASGG---SVTAVTSSVNIIGDLSTINGLNTIISKNGKNTIVL 1286

Qy 1049 OKSKAKTKVRSKRAVSPDPLDQSLFALBALEVIDAPQSEKOR-LAGEAEKORXK 1107
 Db 1281 KGAELDYKXIQP-----GVASAEVIEAKRALEKQDLSPEEETLAKG 1325
 Qy 1108 DILISY--SNSAL-----SELSATYNSMLSVODELRLFPVDAQSAVMTIAOD 1154
 Db 1326 VSAVAFVEPNNALITVNTQNEFTTRPSQVITI-SEGACACSSGDAVACTVAD 1378
 RESULT 13
 PCT-US04-07001-672
 Sequence 672, Application PC/TUS0407001
 GENERAL INFORMATION:
 APPLICANT: BAKALETZ et al.
 TITLE OF INVENTION: GENES OF AN OTTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
 FILE REFERENCE: 28335/38815A
 CURRENT APPLICATION NUMBER: PCT/US04/07001
 CURRENT FILING DATE: 2004-03-05
 PRIOR APPLICATION NUMBER: US 60/453,114
 PRIOR FILING DATE: 2003-03-06
 NUMBER OF SEQ ID NOS: 771
 SOFTWARE: Patentin version 3.2
 SEQ ID NO 672
 LENGTH: 1492
 TYPE: PRT
 ORGANISM: H. influenzae
 PCT-US04-07001-672
 Query Match 3.8%; Score 275; DB 1; Length 1492;
 Best Local Similarity 19.1%; Pred. No. 7.5e-07;
 Matches 274; Conservative 192; Mismatches 484; Indels 484; Gaps 62;
 Qy 42 ENKKEFTVGAQNIKYNNQGOVSTKTAPEMLDVSIVSRGVALVEQYIVSAENVG 101
 Db 219 KNEGVISVNGSISL--AGQKITISDIINPTIYSI-----AAPEKAI-----NLG 264
 Qy 102 YTDVFGAEGNPDQHRFTYKIVKRNRYKKOMLHPYEDDYHNPLHFTVBAPIDMSN 161
 Db 265 ---DIFAKGNIN-----VRAANINOG-----KLADSVSDSGKN 298
 Qy 162 MNGSTTSORTKTPERVRIKSGRGFWANDQKQD--VAGAYHYITAGANTHORGANGYS 219
 Db 239 IYLSAKEGAE-----IG--GVISAQNOQAKGKMLTGDVTLTKGAVIYDLSGREGET 351
 Qy 220 YLGGDYKXAGEYGRPLIAGSGKDSGSPMTYAEQO---KW-----LINGILREKPPG 271
 Db 352 YLGGDERGKNGIOLAKTISLEKSTINVSKEKGRATWGDIALIDGINAOGKDIA 411
 Qy 272 KENGFOUAKSFEDEFERDLHTSLYTRAGNGVTTISGNDGSGITQSGIPSEIKITL 331
 Db 412 KTGGF-----VETS-----GHYLSIGND--AAVEAKEMWLPDQVIT 446
 Qy 332 ANMSLEPKEDKVNPRYDGN-----IYSPILNGETLTFMOKOGSLIFA 378
 Db 447 SNGN---DDQSOLKDRDSDPNKILADNKHVYNNKLTSTALAKGIVNISAKKCVNTAD 503
 Qy 379 SDINOGA-----GGLYFEGNFTVSPNSNOTWQ--GAGIVSNSNV----- 417
 Db 504 INHNGTLTLHSEOGGYEINGJLTSBONAKLTIRAGSWVYVHAKNTITMGFLNITAGGSV 563
 Qy 418 -----TWKNGVEHRLSKIGKGLTHVQAKENGK 447
 Db 564 APERAGDGRASAPAKIYAQVITAGSGODFRNNVS---INGRGGLKFTAKG--NGG 619
 Qy 448 SISVG-----DGKVLIEGQAD-----QNKQAFSEIGLVSGRGTQ 484
 Db 620 NFSKEDGYLANSINISHTANNQSLYFHRQGYTWNLTOLNVDSDSFLTSIXDAIK 679
 Qy 485 IN-----DDQ-----FDTKFEFGGRLDNGSLTFKRIQ-----TD 521
 Db 680 VGGYDAKDKKNTGGIFTRDITFNVKQARVDIS--YTLPISPVKNRSLAANPDGNITV 738

QY 522 EGAMLVN-----HNTQANVTTC-----541
 DB 739 KGGVVLKFNALSNRYKTCGVNISRFINVFGRLNLSMPSTTLEFVANDLIINAT 798
 QY 542 -----NESIVLPGN-----NINKLJYKEIAY-----NGWGE 570
 DB 799 NSFSVKSIEBGTDLHDLTGKYNNAVITIKGNTVLSKNATKEDKRVYVKGALTLAS 858
 QY 571 TQKKNKGRNLNLYKCTTEDRTLLISGTVLKGDIT-----QTKGLFFSGRPYH 621
 DB 859 ANFNHKGALTVAGNINIQGLVATGDTIDVSDGFVGNDAFNGNTNMLNITGNFTNN 918
 QY 622 AVN-----YLNKRSEEGEGPGEIYWHOMKIRTKAFNE 657
 DB 919 GTSIIDYKGAAGKAGNTNBSGLNITTHAN-----TQKTLITGNITNKKGLIRDNKNA 975
 QY 658 QTKGSAAVERNVSSLEGNWTVS-----NNANATFGVWNOQNTICTR- 700
 DB 976 EIQIGG-----NISCKEGLTISSDKXNITKQITIKAGVNGENSDSTENMANLITKTK 1030
 QY 701 -----SWTGLTTCQV-----DITDTKY-----719
 DB 1031 LELTNLNLISGFHKAETITAKONSDLIIGKASSDSGNAGQKVIPEKXOSKISAGNHVT 1090
 QY 720 INSIPKTCJNGSINLTDNATANKVL-----AKNGVTLTNHSOFTLSNNAQTQIR 773
 DB 1091 LMSVEVTS-NGNNSNAGDSNGNAGLTIKADYAVANNIT-----SHKTNISAT-TGNTV 1144
 QY 774 LSNST-----ATVDNANLNGVHLTDSAQFSLNHSFHQIGSKGT-----VTLEN 822
 DB 1145 TKEGTINATTVGVEVYAKTGIDKIGIE-SKSGVYTLNATCDL-AVGNISGNTVSTANS 1203
 QY 823 ATWMPSTTLQKTLNLSSTITLMSAYSASNNTPRRSLTETPTTSABHEFTLLYNG 882
 DB 1204 GLTVKADSTIKG-TGSVTLISQSGDIGTISGKTYSV-TATLT-----DGLTVKG 1251
 QY 883 --KLISG--QGTFOETSSLFPGYKSDKLKLSNDASGDYILSVR--NNGKE-----PETHQ 930
 DB 1252 GAKINATEGTAITLASSSKLTTEASSSTISAKQVDLSAKGNGGSIANAANVTLLTGT 1311
 QY 931 LTVESKDNOPJSDKIKFTLNDHVDAGALRYKLVNDGEFRLNHPKQELHNDLVRAE 990
 DB 1312 LITVYKSSINANSGLTVINAEKALDGT-----1340
 QY 991 QAEPTLEAKQVECTATQGBPRKRSRRAKAFPTLPDQSLMALE--AKQAEPTTET 1048
 DB 1341 SGRDVT--VANATNASGG-----SVTAVTSSVNTIGDLSITNGMLTISKNGNTVVL 1391
 QY 1049 QKSAKTKKVRSKRAVSPDLQSLFLAALFVLDAPQSEKDR-LAQBARKQKQK 1107
 DB 1392 KGAELIDVYIQP-----GASAEVITEKRLKEKYKDLSDERETLAKLG 1436
 QY 1108 DLISRY--SNSAL--SELATVNSMLSVQDELDFVDOAGSVMTNIAQD 1154
 DB 1437 VSAVEFEYEPNNAITVATQNEFTTSPSSQYTI-SEKACFSSGDAAVETNVAD 1489

RESULT 14

/ Sequence 41, Application US/039806709
 / GENERAL INFORMATION:
 / APPLICANT: LOOSMORE, Sheena M.
 / APPLICANT: YANG, Yan-Ping
 / APPLICANT: KLEIN, Michel H.
 / TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 / FILE REFERENCE: 1038-1138 MIS
 / CURRENT APPLICATION NUMBER: US/09/806,709
 / PRIOR APPLICATION NUMBER: PCT/CA99/00938
 / PRIOR FILING DATE: 1999-10-07
 / PRIOR APPLICATION NUMBER: 09/167,568
 / PRIOR FILING DATE: 1998-10-07

/ PRIOR APPLICATION NUMBER: 09/206,942
 / PRIOR FILING DATE: 1999-12-08
 / NUMBER OF SEQ ID NOS: 95
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 41
 / LENGTH: 1005
 / TYPE: PRT
 / ORGANISM: Haemophilus influenzae
 / US-09-806-709-41

Query Match 3.7%; Score 269; DB 5; Length 1005;

Best Local Similarity 20.5%; Pred. No. 9.5e-07;

Matches 224; Conservative 144; Mismatches 363; Indels 364; Gaps 54;

QY 313 GGGSIYOKSGIPSEIKITLNLNMLPKEDKRVHPRDGFNYSBRLNNGELTFYMOQK 372
 DB 19 GRGDTNRYKSAAGLTSINNAITSLAQGIEV-----NISAKK--NYVWADVYKN 68
 QY 373 GSILFASDINQAGGLYFEGNFTVSPNSQWQAG--IHYV-----412
 DB 69 GLVHLSQEN--GVKINGNITSTONGNLTITKGVKVDVHKRITLGMGFLNITSNNIT 124
 QY 413 -----NSVTVWKNVGEHNR-----LSKIGKGLTVQAKGN-----KGISVGD 454
 DB 125 PEKGDVLTITKQGNINQENKOLRPSVSLMGAGLFTTANKGNHTKEDVTLNI-SG 183
 QY 455 KVLISQ-----QADDQ-----GNKQAFSEIGLYSGGYQLND-DKQPTD 494
 DB 184 KVINQTPTHMLAPWNASADSRYNTVTLTLGNNAQFTPIKFPDSNRVVALNNGSRPAG 243
 QY 495 KEFFGGRGRDLNHSILTKRLQNTD-EGKATVTHNTTQ-----AAVTTTGNES- 544
 DB 244 KFY-----GK-----NNMKENIGDANAYEFLKSDNTSNKRPDLQSLNLSATNGTV 294
 QY 545 -----IVPNGNI-----NKLDRKEIANGWGE-----570
 DB 295 SEDIHANLSASTELNLSINISNVNSHSHRGNAPEIKDLILNAGSFYNLKQ 354
 QY 571 TDKNKNGSLNLVYKPTTEDRTL-----JSG-----GTNLKGDIT-QTKGLFFSGRPYH 621
 DB 355 KDFNDVSEKKAIF--STHNLITLGGVTLGGENSSNIKGNININSKAVNTLQA---H 408
 QY 622 A-TNHLNREWEMEGIPQGEITWDEHINRTPKAEHQKGSAAVSVNVSIEGWTVS 680
 DB 409 AGTSHLDKX-----ERTLLGAVSVGSLNLTGSN-AHIDGSLSLA 448
 QY 681 NNANATFGVWNOQNTICTSDMTGLTTCQKVDLTQKYNISIPYQ-----INGSINTLD 736
 DB 449 ESAKQV-GKINNNTI-----TG--TFNNMGADINIKQGVYLGQDITNNGNINITT 498
 QY 737 NATANYKGLANQNT-----LTNHSOFTLSNNAQT-IGNIRLS-----775
 DB 499 NASVQKTI--INGNITNKKGLNLIKDKANAEIQGNI-SQKEGNLTISGDYINITKRI 556
 QY 776 -----UNSTATVDNANL-----NANVHLTDSQFSLK-----802
 DB 557 EYKADTQGSQSCVSNANLTIKTKELTLTDNLNLSGNKAEITAKONSDLIIGASSD 616
 QY 803 NSHFHSHQIGGX-----GTVTLNATWTPSDTLTQW-----LITNNSITTL 845
 DB 617 NSM-AKQITPDKVMDSKISAGNHVNTLNSKVEITSNDSGTGSGDDNNTIGTISADYTV 675
 QY 846 NS-----AYSASSNTPRRRSLETETPTTSABHEFTLLVNGKLSGQGTQFISLF 897
 DB 676 NSNITSHKTWNISASBQGITTYAGTINATTVGSEVYAKTGDISGTLISG-----724
 QY 898 GYKSDKLSNDARGVITLSEVNTGKEPETLEQVLVESKQNOPLSDKLFTLNDHVA 957
 DB 725 --KIVSVATLDS-----LTVGKAGKINABGTAITLASSG-----KLITFANSALS 769
 QY 958 GAFRYKLVNDGEFRLNHPKEQELHNDLVRAEQAEITLAKQVEPTAK---TQTEPER 1013
 DB 770 GA-----NGVTASSQSGDI-----SGTISGKTYSVTAASSGSLTVGSDAK 808

QY 1014 VSRRAARAP-----DTLPDSLNALAEK-Q 1041
 DB 809 INATEGALITLTKGTLTVKGSNIADMEGLVINAODATLNGDSGRTEVNAVNASGS 868
 QY 1042 AELTATQSKAKT-----KVRSKRAVPSDPLDQSLF-ALAEALVIDAP 1087
 DB 869 GNVTAKTSSSVITGDSLTNGLNITISKNGKVTVLKGAIDVKYIQGVAANEVIEAK 928
 QY 1088 QOSEKDR-LAOEBAKQKOKODLSRY--SNSAL-----SELSATVNSMLSVODELDRLF 1139
 DB 929 FALKVKDLSDEERETLAKLGSVAVFIEPNNITLVTQIEFTTPSSQVTL-SEKACF 987
 QY 1140 VDAQASAVMTNIAOD 1154
 DB 988 SSGNGAAVCTNVAAD 1002

RESULT 15
 US-09-806-709-39
 ; Sequence 39, Application US/09806709
 ; GENERAL INFORMATION:
 ; APPLICANT: LOOSMORE, Sheera K.
 ; APPLICANT: YANG, Yan-Ping
 ; APPLICANT: KLEIN, Michel H.
 ; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
 ; FILE REFERENCE: 1038-1118 MTS
 ; CURRENT APPLICATION NUMBER: US/09/806,709
 ; CURRENT FILING DATE: 2001-04-04
 ; PRIOR APPLICATION NUMBER: PCT/CA99/00938
 ; PRIOR FILING DATE: 1999-10-07
 ; PRIOR APPLICATION NUMBER: 09/167,568
 ; PRIOR FILING DATE: 1998-10-07
 ; PRIOR APPLICATION NUMBER: 09/206,942
 ; PRIOR FILING DATE: 1999-12-08
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 39
 ; LENGTH: 1011
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-09-806-709-39

Query Match 3.7%; Score 269; DB 5; Length 1011;
 Best Local Similarity 20.5%; Pred. No. 9.6e-07;
 Matches 224; Conservative 144; Mismatches 363; Indels 364; Gaps 54;

QY 313 GQGSITQKSGIPSEIKITLANKSLPLEKDKVHNPYDGNIVSPRLNSETLYPMQKQ 372
 DB 25 GRGDPNKVSAEGLTSINNAKLSTALQKGLV-----NISATK--WTVRAADVAKN 74
 QY 373 SGLIADINDQAGCYLFGNFTVSPNSNTOGAG--HVS-----412
 DB 75 GTLVASQPN---GVKINGNITSYQKGNLITIGKADYKHNTITLGMGLNITSNNIT 130
 QY 413 ---ENSTVTKVNGVEHDR-----LSKIGKGLHQAQNGN-----KGSISVDG 454
 DB 131 FEKENLITTAQGNITSNQENKQRFPSNVLNGMGAGLFTTANKGNTHKFDGTLNI-SG 189
 QY 455 KVLLEQ-----QADDQ-----GNKQAFSEIGLVSGRGTQJLND-DKQPTD 494
 DB 190 KVINQTTPHNIAFWNASADSYWVNTLTLGNNAGFTFIKEFYDMSVALNNGSSRSPAGV 249
 QY 495 KFYGFGRGLDINGSLTEKRIQNTD-EGAMTVHNHTQ-----AANYTTGNES- 544
 DB 250 KFY-----GK---NNEKFNIGDNAYVEFKLSNDNTSNKPLPIQFLSNISATGNQTV 300
 QY 545 -----IVLPNGNI-----NKLDYREKLANVWGFE-----570
 DB 301 SPDIHANSRSTELMMSLNIINSGVFNSHVRGNNAPFEEKDLINATGSPFNKQOT 360
 QY 571 TDKRKHGRLNLIYKPTTEPRTL-----LSG-----GTNLKGLIT-QTKGKLPFGSGRPTPH 621

DB 361 KQKFDNSTERKALF--STHNLITIGSVTLGGNSSSNKKNINIKNAVTLQA-----H 414
 QY 622 A-YNHLNRSSEWEKIPQGEIWDHWINRTFAENFQIKGSAAVSRNWSIEGNWTVS 680
 DB 415 AGTSLDKK-----ERTLIGNSVGNLNIISN-AHIDNLSIA 454
 QY 681 NNANATFGVVPNOQNTICTRSDWTGLTTCQKVDLDTQVINSIPKQ-----INGSLNLT 736
 DB 455 BSARKQ-GKTNNNLNI-----TG--TFTNGTADINIKQGVVKLGQDITNNGNLNITT 504
 QY 737 NNTANVGLAKANGNT-----LTHSQFTLSNNAQ-IGNRILS-----775
 DB 505 NASTVQKTI--INGNITKKGDLNIXDIKANMELIQENIQKQGNLITLSDKINITYRI 562
 QY 776 -----DNSTATVDNANL-----NGNHLTDSAQPSLK-----802
 DB 563 EIXADTQGNSDGVAANALITIKRELITLTDNLNINSGPNAELITAKNSDILLIGKASD 622
 QY 803 NSHFSHOIQGDK-----GTTVLNATWTMPSDITLQN-----LTLNNSITTL 845
 DB 623 NSN-AXQITFDKVKOSKISAGNHVTLNSKVEFNSDGTGKSDDNNGLITISADTV 681
 QY 846 NS-----AVSASNNTPRRRLSEFTTTPTSAREHENTLTVNGKLSGQGTPTQSSLF 897
 DB 682 NSNITSKRTVNISASGGITTKAGTTIMATGSAVEVTAKTADISGITSG-----730
 QY 898 GKSDXKLKSNDAAGDYILISVNTGKEPELILQTLTVESKDNQPLSDKXKFTLNDHYDA 957
 DB 731 -KIVSVLATDGS-----LVKGAKINITEGATLTLSSG-----KLTEANSALIS 775
 QY 958 GALKYKLVKQDEFFLHPPIKEQELANDVRAEQERTLEAKQVEPTAK-----TOTGEPK 1013
 DB 776 GA-----NGVTASSQSGDI-----SGTISGKTIVSYASSGSLTVGQDAK 814
 QY 1014 VSRRAARAP-----DTLPDSLNALAEK-Q 1041
 DB 815 INATEGALITLTKGTLTVKGSNIADMEGLVINAODATLNGDSGRTEVNAVNASGS 874
 QY 1042 AELTATQSKAKT-----KVRSKRAVPSDPLDQSLF-ALAEALVIDAP 1087
 DB 875 GNVTAKTSSSVITGDSLTNGLNITISKNGKVTVLKGAIDVKYIQGVAANEVIEAK 934
 QY 1088 QOSEKDR-LAOEBAKQKOKODLSRY--SNSAL-----SELSATVNSMLSVODELDRLF 1139
 DB 935 FALKVKDLSDEERETLAKLGSVAVFIEPNNITLVTQIEFTTPSSQVTL-SEKACF 993
 QY 1140 VDAQASAVMTNIAOD 1154
 DB 994 SSGNGAAVCTNVAAD 1002

Search completed: June 23, 2004, 17:37:52
 Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2004, 17:29:18 ; Search time 26 Seconds

(Without alignments)
2767,946 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273

Sequence: 1 MKTVPRLNLTACISLGIV.....SGSGQLGQGVKGLGRW 1394

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7273	100.0	1394	US-08-296-791-2	Sequence 2, Appli
2	7273	100.0	1394	US-09-839-996-2	Sequence 2, Appli
3	7273	100.0	1394	US-10-080-505-2	Sequence 2, Appli
4	7273	100.0	1394	PCT-US95-10661A-2	Sequence 2, Appli
5	7246.5	99.6	1395	US-10-080-505-7	Sequence 7, Appli
6	6115	84.1	1434	US-10-080-505-9	Sequence 9, Appli
7	5827.5	80.1	1411	US-10-080-505-17	Sequence 17, Appli
8	5807	79.8	1436	US-10-080-505-13	Sequence 13, Appli
9	5303.5	72.8	1391	US-10-080-505-15	Sequence 15, Appli
10	5297.5	72.8	1391	US-10-080-505-11	Sequence 11, Appli
11	1673.5	23.0	1541	US-08-296-791-3	Sequence 3, Appli
12	1673.5	23.0	1541	US-09-839-996-3	Sequence 3, Appli
13	1673.5	23.0	1541	US-10-080-505-3	Sequence 3, Appli
14	1673.5	23.0	1541	PCT-US95-10661A-3	Sequence 3, Appli
15	1662.5	22.9	1545	US-08-296-791-4	Sequence 4, Appli
16	1662.5	22.9	1545	US-09-839-996-4	Sequence 4, Appli
17	1662.5	22.9	1545	US-10-080-505-4	Sequence 4, Appli
18	1662.5	22.9	1545	PCT-US95-10661A-4	Sequence 4, Appli
19	1605	22.1	1702	US-08-296-791-5	Sequence 5, Appli
20	1605	22.1	1702	US-09-839-996-5	Sequence 5, Appli
21	1605	22.1	1702	US-10-080-505-5	Sequence 5, Appli
22	1605	22.1	1702	PCT-US95-10661A-5	Sequence 5, Appli
23	1599.5	21.6	1507	US-08-296-791-6	Sequence 6, Appli
24	1599.5	21.6	1507	US-09-839-996-6	Sequence 6, Appli
25	1551	21.3	1848	US-10-080-505-6	Sequence 6, Appli
26	1551	21.3	1848	US-09-839-996-6	Sequence 6, Appli
27	1551	21.3	1848	PCT-US95-10661A-6	Sequence 6, Appli

28	346.5	4.8	73	4	US-10-080-505-44	Sequence 44, Appli
29	344.5	4.7	1569	4	US-09-711-164-312	Sequence 312, App
30	286.5	3.9	1536	1	US-08-038-682-2	Sequence 2, Appli
31	286.5	3.9	1536	1	US-08-302-832-2	Sequence 2, Appli
32	286.5	3.9	1536	2	US-08-530-198-2	Sequence 2, Appli
33	286.5	3.9	1536	2	US-08-469-880-2	Sequence 2, Appli
34	286.5	3.9	1536	2	US-08-728-470-2	Sequence 2, Appli
35	286.5	3.9	1536	2	US-08-617-697-2	Sequence 2, Appli
36	286.5	3.9	1536	3	US-08-719-641-2	Sequence 3, Appli
37	285.5	3.9	1004	4	US-09-206-942-57	Sequence 57, Appli
38	285.5	3.9	1010	4	US-09-206-942-55	Sequence 55, Appli
39	283.5	3.9	1222	4	US-09-206-942-37	Sequence 37, Appli
40	283.5	3.9	1228	4	US-09-206-942-34	Sequence 34, Appli
41	283.5	3.9	1536	4	US-09-206-942-67	Sequence 67, Appli
42	283	3.9	1477	1	US-08-038-682-4	Sequence 4, Appli
43	283	3.9	1477	1	US-08-302-832-4	Sequence 4, Appli
44	283	3.9	1477	2	US-08-530-198-4	Sequence 4, Appli
45	283	3.9	1477	2	US-08-469-880-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-296-791-2
Sequence 2, Application US/08296791
Patent No. 6245337

GENERAL INFORMATION:

APPLICANT: St. Geme III, Joseph W.

TITLE OF INVENTION: Haemophilus Adherence and Penetration

TITLE OF INVENTION: Protein

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSER: Flehr, Hobach, Test, Albritton & Herbert

STREET: 4 Embarradero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: United States

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/296,791

FILING DATE: 25-AUG-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Treccartin, Richard F.

REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLCULE TYPE: protein

Query Match 100.0%; Score 7273; DB 3; Length 1394;

Best local similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 1394; Conservative 0; Indels 0;

QY 1 MKTVPRLNLTACISLGIVSQAMAGHTTFTGIDYVDFENKGRFTVGAQNIKVYVKQ 60
DB 1 MKTVPRLNLTACISLGIVSQAMAGHTTFTGIDYVDFENKGRFTVGAQNIKVYVKQ 60

```

QY 61 GQVGTSTMTAEMIDPSSVSENGVAAJVENQYTVSAHVNGTVDVDEGAGNPNPQHRFT 122
DB 61 GQVGTSTMTAEMIDPSSVSENGVAAJVENQYTVSAHVNGTVDVDEGAGNPNPQHRFT 122
QY 121 YKIVKRNKYNKKNLHEDDYDYNPLKFTTEAPIDMTSNNGSTYSRTRYPERVIG 186
DB 121 YKIVKRNKYNKKNLHEDDYDYNPLKFTTEAPIDMTSNNGSTYSRTRYPERVIG 186
QY 181 SGRQFENDDQKGDQVAGAYHLLTAGNTHQORQAGNGSYLAGDVRAGGYPLTJASGR 240
DB 181 SGRQFENDDQKGDQVAGAYHLLTAGNTHQORQAGNGSYLAGDVRAGGYPLTJASGR 240
QY 241 GDSGSPWETIDAEKQKMLINGILREGNPFEGKNGEQVLRKSYDELFERDLHTSLYTRA 300
DB 241 GDSGSPWETIDAEKQKMLINGILREGNPFEGKNGEQVLRKSYDELFERDLHTSLYTRA 300
QY 301 GNGVYTTISGNDNGQSGITQKSGIPSEIKITLNMSLPLKEXDKVHNRDGNVYSPRLN 360
DB 301 GNGVYTTISGNDNGQSGITQKSGIPSEIKITLNMSLPLKEXDKVHNRDGNVYSPRLN 360
QY 361 NGETIYEMDQKQSLIFASDINOGAGLYEGNFTVSPNSNQTMQAGIHVSENSTYTMK 420
DB 361 NGETIYEMDQKQSLIFASDINOGAGLYEGNFTVSPNSNQTMQAGIHVSENSTYTMK 420
QY 421 VNGVEHRLSKTIGKGLHVOAKGENKSGISVGDQKXVILEQDADQGNKQAFSEIGLVSGR 480
DB 421 VNGVEHRLSKTIGKGLHVOAKGENKSGISVGDQKXVILEQDADQGNKQAFSEIGLVSGR 480
QY 481 GTVQANDXQFTDTRKYRFGGRGLDJNGSLTFPKRIQNTDBGAMVNHNTQOANVTIT 540
DB 481 GTVQANDXQFTDTRKYRFGGRGLDJNGSLTFPKRIQNTDBGAMVNHNTQOANVTIT 540
QY 541 GNEIYVLPNGNNINKLDYKEJAYNGWEGTDBKNKNGRLNLYKPTEDRILLSGGNT 600
DB 541 GNEIYVLPNGNNINKLDYKEJAYNGWEGTDBKNKNGRLNLYKPTEDRILLSGGNT 600
QY 601 LKGDITQTKGKLPFSSGRPTPHAYNHLNKRWSEMEGIPOGEIYVMDHDMINRTEKAEFQIX 660
DB 601 LKGDITQTKGKLPFSSGRPTPHAYNHLNKRWSEMEGIPOGEIYVMDHDMINRTEKAEFQIX 660
QY 661 GGSVAVSRNVSIEAKWYVSNNAATPGVVPQONTICTRSDWTGLTTOQKVDLDTYVI 720
DB 661 GGSVAVSRNVSIEAKWYVSNNAATPGVVPQONTICTRSDWTGLTTOQKVDLDTYVI 720
QY 721 NSIPTQJNGSINLJDNATAYWKGALNGNTLTFHSGFTLSNNAFOGINTRSDNSTA 780
DB 721 NSIPTQJNGSINLJDNATAYWKGALNGNTLTFHSGFTLSNNAFOGINTRSDNSTA 780
QY 781 TVDNANLNGNVHLTDSAQPSLKNSHFSHQIQDKGTIVYLENATWTPSPDITLQVLTENN 840
DB 781 TVDNANLNGNVHLTDSAQPSLKNSHFSHQIQDKGTIVYLENATWTPSPDITLQVLTENN 840
QY 841 STILINSAVGSASNTPPRRSLTETPTTSAHRNNTLYWGXISQOQFTOTSLFQYK 900
DB 841 STILINSAVGSASNTPPRRSLTETPTTSAHRNNTLYWGXISQOQFTOTSLFQYK 900
QY 901 SDKXLSNDAESDYILSVKNTGKEPETLEQLTLVESKNOPLSDKIKETLNDHVDGAL 960
DB 901 SDKXLSNDAESDYILSVKNTGKEPETLEQLTLVESKNOPLSDKIKETLNDHVDGAL 960
QY 961 RYKLVKNDGEFRLNPIKEQELHNDIVRAEQERTLEAKQYEPFLACTYGEKYSRRAA 1020
DB 961 RYKLVKNDGEFRLNPIKEQELHNDIVRAEQERTLEAKQYEPFLACTYGEKYSRRAA 1020
QY 1021 RAAPEDTLPDQSLNALAEAKOELTETQSAKTKKYSRAVSPDPLDQSLFALEAA 1080
DB 1021 RAAPEDTLPDQSLNALAEAKOELTETQSAKTKKYSRAVSPDPLDQSLFALEAA 1080
QY 1081 LEVTDAPQSEKORLAQEBEAKOKKQKDLISYSNSALSELSTATVSMSTVODELDTLFF 1140
DB 1081 LEVTDAPQSEKORLAQEBEAKOKKQKDLISYSNSALSELSTATVSMSTVODELDTLFF 1140
QY 1141 DQAGSAVNTINIAQDKGRIDSDAPRAYQOQKTNLQIGVQALANGKIGAVSHSRSDNTF 1200

```

```

DB 1141 DQAGSAVNTINIAQDKGRIDSDAPRAYQOQKTNLQIGVQALANGKIGAVSHSRSDNTF 1200
QY 1201 DEQVKNHATLTMGSGFAQYQMGDLQGVNVGTSIGSASKMAEBSKRIHRKAINGVNASY 1260
DB 1201 DEQVKNHATLTMGSGFAQYQMGDLQGVNVGTSIGSASKMAEBSKRIHRKAINGVNASY 1260
QY 1261 QFRLGQLGQIFPGVNNYFERENTQSERVRYKTPSLAFRRNAGIRVDYFTFTDNISY 1320
DB 1261 QFRLGQLGQIFPGVNNYFERENTQSERVRYKTPSLAFRRNAGIRVDYFTFTDNISY 1320
QY 1321 KPYEFVNYDVSNANVQITVNLTVLQOPGRGYMCKEYGLKASILHPOJSAFISQSSGSL 1380
DB 1321 KPYEFVNYDVSNANVQITVNLTVLQOPGRGYMCKEYGLKASILHPOJSAFISQSSGSL 1380
QY 1381 GKOQNVGYLGYRW 1394
DB 1381 GKOQNVGYLGYRW 1394

RESULT 2
US-09-839-996-2
; Sequence 2, Application US/0983996
; Patent No. 6642371
; GENERAL INFORMATION:
; APPLICANT: St. Gene III, Joseph W.
; Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Flehr, Hobbach, Test, Albritton & Hebert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/839,996
; FILING DATE: 20-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-839-996-2

Query Match 100.0%; Score 7273; DB 4; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKKTVFRINFLINACISLGIIVSQAMAGHTYFGIDYQYRDPANNKQKFTVGAQNIKYNNKQ 60
DB 1 MKKTVFRINFLINACISLGIIVSQAMAGHTYFGIDYQYRDPANNKQKFTVGAQNIKYNNKQ 60

```

QY 61 GOLVGTSMKAPMIDFVSVRNGVAALENQYTVSAHNGYTDVDFGAEGNNDQHRFT 120
 DB 61 GOLVGTSMKAPMIDFVSVRNGVAALENQYTVSAHNGYTDVDFGAEGNNDQHRFT 120
 QY 121 YKIVKANNYKXONLHPYEDDYHNPRLKFTYLAAPIDMTSNMGSTYSDBTKYPERVRIG 180
 DB 121 YKIVKANNYKXONLHPYEDDYHNPRLKFTYLAAPIDMTSNMGSTYSDBTKYPERVRIG 180
 QY 181 SGROFMRNDQDKDQVAGAYHYLTAGNTHNQAGANGSYLGGVRAAGSYGPEPLIAGSK 240
 DB 181 SGROFMRNDQDKDQVAGAYHYLTAGNTHNQAGANGSYLGGVRAAGSYGPEPLIAGSK 240
 QY 241 GDSGPMFTYDAEKOKMLINGILREGNPEFGKNGFOLVRKSYDEIPEBDLHTSLYTRA 300
 DB 241 GDSGPMFTYDAEKOKMLINGILREGNPEFGKNGFOLVRKSYDEIPEBDLHTSLYTRA 300
 QY 301 GNGVYTTISGNDNGQSLITQSGIFSEIKITLANSILPLKEXDKVHNPRYDGPNTYSPRLN 360
 DB 301 GNGVYTTISGNDNGQSLITQSGIFSEIKITLANSILPLKEXDKVHNPRYDGPNTYSPRLN 360
 QY 361 NGETLYFMDQKQSLIPASDINOAGGLYFEGNFTVS PMSNQTWQAGIHVSENSITWTWK 420
 DB 361 NGETLYFMDQKQSLIPASDINOAGGLYFEGNFTVS PMSNQTWQAGIHVSENSITWTWK 420
 QY 421 VNGVEHDLSTKIGKTLHQAKGKNGKSIYVGDQKVLBQADQGNKQAPSEIGLVSCR 480
 DB 421 VNGVEHDLSTKIGKTLHQAKGKNGKSIYVGDQKVLBQADQGNKQAPSEIGLVSCR 480
 QY 481 GTVQANDDKQPDQKFEYFEGRGRDLNGHSLFRTIONTDEGAMVHNTTQAAVTT 540
 DB 481 GTVQANDDKQPDQKFEYFEGRGRDLNGHSLFRTIONTDEGAMVHNTTQAAVTT 540
 QY 541 GNSIYVJGNNINKLDYRKEIAVNGMPESTDYKXKHGRNLIVKPTEDDTLLSGGTM 600
 DB 541 GNSIYVJGNNINKLDYRKEIAVNGMPESTDYKXKHGRNLIVKPTEDDTLLSGGTM 600
 QY 601 LKGDITQTKGKLFPSGRPTPHAVNHKMSMEGIPQGEIWDHDMINRTKAEFQIK 660
 DB 601 LKGDITQTKGKLFPSGRPTPHAVNHKMSMEGIPQGEIWDHDMINRTKAEFQIK 660
 QY 661 GGSVSVRNVSSJEGMWTYSNNANATFGVPPNQNTTCTRSDMTGLTQCKXDLDPRTVI 720
 DB 661 GGSVSVRNVSSJEGMWTYSNNANATFGVPPNQNTTCTRSDMTGLTQCKXDLDPRTVI 720
 QY 721 NSIPKTOINSIMLTNATANYKGLAKINGNTLTMSOFTLSNNATQIGNRLSDNSTA 780
 DB 721 NSIPKTOINSIMLTNATANYKGLAKINGNTLTMSOFTLSNNATQIGNRLSDNSTA 780
 QY 781 TYDNANLNGNVHLTDSNOFSLKNSHFSHOIQGKGTTLLENAITMPSDTTLQNLTLAN 840
 DB 781 TYDNANLNGNVHLTDSNOFSLKNSHFSHOIQGKGTTLLENAITMPSDTTLQNLTLAN 840
 QY 841 STTLNSAYSASNNTPRRSLFETITTSABERFNTLTNGKLSGGOTQFTSSJFGYK 900
 DB 841 STTLNSAYSASNNTPRRSLFETITTSABERFNTLTNGKLSGGOTQFTSSJFGYK 900
 QY 901 SDKLSKNSABGDYLLSVNTGKEPTELEQTLVESKONPLSDKLTLENDHVDAGAL 960
 DB 901 SDKLSKNSABGDYLLSVNTGKEPTELEQTLVESKONPLSDKLTLENDHVDAGAL 960
 QY 961 RYKLVKNDGEPFLNHPKIOELHNDVRAEQEFTLEAKOVEPLATQTPGPKRSRRA 1020
 DB 961 RYKLVKNDGEPFLNHPKIOELHNDVRAEQEFTLEAKOVEPLATQTPGPKRSRRA 1020
 QY 1021 RAAPPTLPDQSLNALBAKQABLTETQSKAKTKKRSKRAVSDPLDQSLFALBA 1080
 DB 1021 RAAPPTLPDQSLNALBAKQABLTETQSKAKTKKRSKRAVSDPLDQSLFALBA 1080
 QY 1081 LEVIDAQOQSEKORLAQEPKQKQKOLISYNSALSSELSATVSMLSYQDELDRLFV 1140
 DB 1081 LEVIDAQOQSEKORLAQEPKQKQKQKOLISYNSALSSELSATVSMLSYQDELDRLFV 1140

QY 1141 DQAGAWTMIADPKRRYDSAPRAYOQKTNLBOIGYOKALANGRIGAVFSSHSNDTF 1200
 DB 1141 DQAGAWTMIADPKRRYDSAPRAYOQKTNLBOIGYOKALANGRIGAVFSSHSNDTF 1200
 QY 1201 DBOVGNHATLMMSGFAQYQWMDQPGVNVGTISI SASMAEBSQSKHRAINYGNASY 1260
 DB 1201 DBOVGNHATLMMSGFAQYQWMDQPGVNVGTISI SASMAEBSQSKHRAINYGNASY 1260
 QY 1261 QPRLGOLGQYFQVNRYPLEENYQSEBRVYKTPSLANRNACIRVDYFTPTNINISV 1320
 DB 1261 QPRLGOLGQYFQVNRYPLEENYQSEBRVYKTPSLANRNACIRVDYFTPTNINISV 1320
 QY 1321 KPYEFNVYDVSQANVOTTNLTVLQOPFGRYMKEVGLAKABILHROISAFISQSGSL 1380
 DB 1321 KPYEFNVYDVSQANVOTTNLTVLQOPFGRYMKEVGLAKABILHROISAFISQSGSL 1380
 QY 1381 GKQOYNGVGLGYRW 1394
 DB 1381 GKQOYNGVGLGYRW 1394

RESULT 3
 US-10-080-505-2
 ; Sequence 2, Application US/10080505
 ; Patent No. 6676948
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RET/DCF/DHR
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 1394
 ; TYPE: PR
 ; ORGANISM: Haemophilus influenzae
 ; US-10-080-505-2

Query Match 100.0%; Score 7273; DB 4; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTTFRLNPLFACISIGIVSOAMAGHTYFGIDYQYTRDFAENKKEFTVGAQNTKYVKNQ 60
 DB 1 MKTTFRLNPLFACISIGIVSOAMAGHTYFGIDYQYTRDFAENKKEFTVGAQNTKYVKNQ 60
 QY 61 GOLVGTSMKAPMIDFVSVRNGVAALENQYTVSAHNGYTDVDFGAEGNNDQHRFT 120
 DB 61 GOLVGTSMKAPMIDFVSVRNGVAALENQYTVSAHNGYTDVDFGAEGNNDQHRFT 120
 QY 121 YKIVKANNYKXONLHPYEDDYHNPRLKFTYLAAPIDMTSNMGSTYSDBTKYPERVRIG 180
 DB 121 YKIVKANNYKXONLHPYEDDYHNPRLKFTYLAAPIDMTSNMGSTYSDBTKYPERVRIG 180
 QY 181 SGROFMRNDQDKDQVAGAYHYLTAGNTHNQAGANGSYLGGVRAAGSYGPEPLIAGSK 240
 DB 181 SGROFMRNDQDKDQVAGAYHYLTAGNTHNQAGANGSYLGGVRAAGSYGPEPLIAGSK 240
 QY 241 GDSGPMFTYDAEKOKMLINGILREGNPEFGKNGFOLVRKSYDEIPEBDLHTSLYTRA 300
 DB 241 GDSGPMFTYDAEKOKMLINGILREGNPEFGKNGFOLVRKSYDEIPEBDLHTSLYTRA 300
 QY 301 GNGVYTTISGNDNGQSLITQSGIFSEIKITLANSILPLKEXDKVHNPRYDGPNTYSPRLN 360
 DB 301 GNGVYTTISGNDNGQSLITQSGIFSEIKITLANSILPLKEXDKVHNPRYDGPNTYSPRLN 360
 QY 361 NGETLYFMDQKQSLIPASDINOAGGLYFEGNFTVS PMSNQTWQAGIHVSENSITWTWK 420
 DB 361 NGETLYFMDQKQSLIPASDINOAGGLYFEGNFTVS PMSNQTWQAGIHVSENSITWTWK 420

Db 361 NEEITYEMDQKQSLIFASDINQAGGLYFEGNFTVSPNSQTVQAGAHVSENSTVTWK 420
 QY 421 VNGVEHDLRSKLGKTLHQVAKGENKKSISVGGCKYTLTCCQADQKQKQAFSTIGVSR 480
 Db 421 VNGVEHDLRSKLGKTLHQVAKGENKKSISVGGCKYTLTCCQADQKQKQAFSTIGVSR 480
 QY 481 GTVQJNDKQKQFDKFEYFGFRGRLDLNHSGLTEKRIQNTDEGAMTVNHTTQAAVTT 540
 Db 481 GTVQJNDKQKQFDKFEYFGFRGRLDLNHSGLTEKRIQNTDEGAMTVNHTTQAAVTT 540
 QY 541 GNESTIVLNGNLYKLDYRKETIAYNGFGBFDKNGKNGRLNIITYPTEDRTLLSGGTN 600
 Db 541 GNESTIVLNGNLYKLDYRKETIAYNGFGBFDKNGKNGRLNIITYPTEDRTLLSGGTN 600
 QY 601 LKGDITOTKGLFPGSRPTPHAYNHNKNSMEGIPCEIYTHDHTINTKQKNEQIK 660
 Db 601 LKGDITOTKGLFPGSRPTPHAYNHNKNSMEGIPCEIYTHDHTINTKQKNEQIK 660
 QY 661 GGSAAVSNVSSIEGKMTVSNNAATEGVVPMQONTICTRSDPTGLTTCQKVDLTVTKY 720
 Db 661 GGSAAVSNVSSIEGKMTVSNNAATEGVVPMQONTICTRSDPTGLTTCQKVDLTVTKY 720
 QY 721 NSIPTQINSINLTNATANYKGLAKNGVTLTNHSGFTLSNNAIQGNIRLSDNSTA 780
 Db 721 NSIPTQINSINLTNATANYKGLAKNGVTLTNHSGFTLSNNAIQGNIRLSDNSTA 780
 QY 781 TVDNANLNGVHLTDSQPSLKSHSHSQIQGKGTVTLENATWMPSTTLQVLTNN 840
 Db 781 TVDNANLNGVHLTDSQPSLKSHSHSQIQGKGTVTLENATWMPSTTLQVLTNN 840
 QY 841 STTLNSAYSASSNNTPRRSLETTPTSAHERPNTLVNGKLSGGTFCQTSLSFYK 900
 Db 841 STTLNSAYSASSNNTPRRSLETTPTSAHERPNTLVNGKLSGGTFCQTSLSFYK 900
 QY 901 SRIKLSNDAGDYLLSVNNGKPELTLEOLTVESKNOPLSDMLKFTLENDHYDACL 960
 Db 901 SRIKLSNDAGDYLLSVNNGKPELTLEOLTVESKNOPLSDMLKFTLENDHYDACL 960
 QY 961 RYKLVKNDGEFLNPIKECEIHNLDVRAQAEFLTEAKQVEPNAKTQGEKVSRRRA 1020
 Db 961 RYKLVKNDGEFLNPIKECEIHNLDVRAQAEFLTEAKQVEPNAKTQGEKVSRRRA 1020
 QY 1021 RAAPDITLPOQSLNLALEKQAEITAEQSKAKTKVRSRAVFSDDLQSLPALAA 1080
 Db 1021 RAAPDITLPOQSLNLALEKQAEITAEQSKAKTKVRSRAVFSDDLQSLPALAA 1080
 QY 1081 LEVIDAPOQSEKORLAQEBEAKOKKODLISRSNSALSELATVNSMLSVODELDELTV 1140
 Db 1081 LEVIDAPOQSEKORLAQEBEAKOKKODLISRSNSALSELATVNSMLSVODELDELTV 1140
 QY 1141 DOASAVWNTIADKRRYSDAFRAVQCCQKMLRQIGVQKALNGRLGAVESHRSNDTF 1200
 Db 1141 DOASAVWNTIADKRRYSDAFRAVQCCQKMLRQIGVQKALNGRLGAVESHRSNDTF 1200
 QY 1201 DEOVKNEATLTWMSGEAQQWGDLOFGVNVGIGISASMAEBSRKIRKALNTGVNAST 1260
 Db 1201 DEOVKNEATLTWMSGEAQQWGDLOFGVNVGIGISASMAEBSRKIRKALNTGVNAST 1260
 QY 1261 QERIGUQIOPYEGNRRFIRENYSSEBRKXTSLAPNRYNAGTRVDYTFPTDINISY 1320
 Db 1261 QERIGUQIOPYEGNRRFIRENYSSEBRKXTSLAPNRYNAGTRVDYTFPTDINISY 1320
 QY 1321 KPYFVYVYDVSNANAVOTTVALTVLOQFGRWCKEVLKAEILHFQISAFISQSSQL 1380
 Db 1321 KPYFVYVYDVSNANAVOTTVALTVLOQFGRWCKEVLKAEILHFQISAFISQSSQL 1380
 QY 1381 GKQONVGVKLGIRN 1394
 Db 1381 GKQONVGVKLGIRN 1394

RESULT 4
 PCT-US95-10661A-2

; Sequence 2, Application PC/TUS9510661A
 ; GENERAL INFORMATION:
 ; APPLICANT: Washington University, et al.
 ; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Piehr, Hobbach, Teet, Albritton & Herbert
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/10661A
 ; FILING DATE: 16-AUG-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/296,791
 ; FILING DATE: 25-AUG-1994
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Treacartin, Richard F.
 ; REGISTRATION NUMBER: 31,801
 ; REFERENCE/DOCKET NUMBER: FP-59941/RFT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 396-3249
 ; TELE: 9-0 277299
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1394 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US95-10661A-2
 Query Match 100.0%; Score 7273; DB 5; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKTVERLNFLACISLIGVSCAMAGHTYFGIDYQYRFAENKQFTVGAQNTKYNQ 60
 Db 1 MKTVERLNFLACISLIGVSCAMAGHTYFGIDYQYRFAENKQFTVGAQNTKYNQ 60
 QY 61 GQIVGSMTKAPMTDFSVSRNGVAALVENQYIVSAHVNGYTVDFGAEQNPDDHRET 120
 Db 61 GQIVGSMTKAPMTDFSVSRNGVAALVENQYIVSAHVNGYTVDFGAEQNPDDHRET 120
 QY 121 YKIVKNNYKQNDLHPYEDYHNPRLHGFVTEAPIDMTSNNGSTYSDRTKXPERVIG 180
 Db 121 YKIVKNNYKQNDLHPYEDYHNPRLHGFVTEAPIDMTSNNGSTYSDRTKXPERVIG 180
 QY 181 SGRQWRNDQKGDQVAGAYHYLTAGNTHNGRQAGNGYSYLGQDVRAKAGEYGLTAGSK 240
 Db 181 SGRQWRNDQKGDQVAGAYHYLTAGNTHNGRQAGNGYSYLGQDVRAKAGEYGLTAGSK 240
 QY 241 GSGSGPMFTYAEKQKMLNGILAREGNPESEKNGQVYRKSYDEIPEQDLSITYTA 300
 Db 241 GSGSGPMFTYAEKQKMLNGILAREGNPESEKNGQVYRKSYDEIPEQDLSITYTA 300
 QY 301 GNGVYVYVYDVSNANAVOTTVALTVLOQFGRWCKEVLKAEILHFQISAFISQSSQL 360
 Db 301 GNGVYVYVYDVSNANAVOTTVALTVLOQFGRWCKEVLKAEILHFQISAFISQSSQL 360
 QY 361 NEEITYEMDQKQSLIFASDINQAGGLYFEGNFTVSPNSQTVQAGAHVSENSTVTWK 420
 Db 361 NEEITYEMDQKQSLIFASDINQAGGLYFEGNFTVSPNSQTVQAGAHVSENSTVTWK 420

QY 421 VNGVEHRLSKIGKGLTHVQAKGKNGKSIISVGDGVILLEQADDOQKQAFSEIGLVSGR 480
 DB 421 VNGVEHRLSKIGKGLTHVQAKGKNGKSIISVGDGVILLEQADDOQKQAFSEIGLVSGR 480
 QY 481 GTVQLNDCKQFDTDKFFGFRGRLDLNGHSLTFKRIQNTDEGAMTVNHTTOAANVTIT 540
 DB 481 GTVQLNDCKQFDTDKFFGFRGRLDLNGHSLTFKRIQNTDEGAMTVNHTTOAANVTIT 540
 QY 541 GNEISIVLPNGNNINKLDRKEIAVNGFETDKNHNRNLTYKPTEDRLLLSGNTN 600
 DB 541 GNEISIVLPNGNNINKLDRKEIAVNGFETDKNHNRNLTYKPTEDRLLLSGNTN 600
 QY 601 LKGDITQTKKLFPSGRPTPHAVNHLNKRSEMEGIPGGEIVWDHWINRTKAKENFOIK 660
 DB 601 LKGDITQTKKLFPSGRPTPHAVNHLNKRSEMEGIPGGEIVWDHWINRTKAKENFOIK 660
 QY 661 GGSAYVSRVSSIEGNMTVSNNAATFGVVPNOQNTICTSDMTGLTTCQKVDLDTKVI 720
 DB 661 GGSAYVSRVSSIEGNMTVSNNAATFGVVPNOQNTICTSDMTGLTTCQKVDLDTKVI 720
 QY 721 NSIPKTOINGSINTLDNATANVGLAKNGVNTLTNHSOFTLSNNAVOIGNIRLSDNSTA 780
 DB 721 NSIPKTOINGSINTLDNATANVGLAKNGVNTLTNHSOFTLSNNAVOIGNIRLSDNSTA 780
 QY 781 TVDNAMANGNVELTDSAQFSJKNSESHQIQDGKGTVILENATWTMSDITLQNTLNN 840
 DB 781 TVDNAMANGNVELTDSAQFSJKNSESHQIQDGKGTVILENATWTMSDITLQNTLNN 840
 QY 841 SITLINSAYSASNNTPRRSLTETPTPSAHRNTLYNGCLSGOGTFPFTSLFGYK 900
 DB 841 SITLINSAYSASNNTPRRSLTETPTPSAHRNTLYNGCLSGOGTFPFTSLFGYK 900
 QY 901 SDKLKSNDAEGDYLSVNTGKEPTELEOLIVESKONPDSDKLKTLENDHVNAGL 960
 DB 901 SDKLKSNDAEGDYLSVNTGKEPTELEOLIVESKONPDSDKLKTLENDHVNAGL 960
 QY 961 RYKLVKNDGEFRLNPIKEQELANDLVRAEQERTLEAKQVETLAKTQTEPKYRSRRA 1020
 DB 961 RYKLVKNDGEFRLNPIKEQELANDLVRAEQERTLEAKQVETLAKTQTEPKYRSRRA 1020
 QY 1021 RAAPFTLEPDSLLNALAEQAQALTFQKSKATKTVSKRAVPSDLDOSLFALERA 1080
 DB 1021 RAAPFTLEPDSLLNALAEQAQALTFQKSKATKTVSKRAVPSDLDOSLFALERA 1080
 QY 1081 LEVYIDAPQOSSEKORLAQEBAROKOKULISRYNSALSLSATNWSLSTVOEELRLV 1140
 DB 1081 LEVYIDAPQOSSEKORLAQEBAROKOKULISRYNSALSLSATNWSLSTVOEELRLV 1140
 QY 1141 DOQSAVMTNIAJDKRRYDSDAFRAYOQOKTNLQIGVQKALANRIGAVFSHSRSDNTF 1200
 DB 1141 DOQSAVMTNIAJDKRRYDSDAFRAYOQOKTNLQIGVQKALANRIGAVFSHSRSDNTF 1200
 QY 1201 DEGVKNHATLTMGSGFYQYQWMDLQFVNVVGTGIGSASKRAEQSKTHRAKINYNVASY 1260
 DB 1201 DEGVKNHATLTMGSGFYQYQWMDLQFVNVVGTGIGSASKRAEQSKTHRAKINYNVASY 1260
 QY 1261 QFRLQGLGLOPFEGNVYFLEKRNQSEBVAVKPPLAENRNASIRDTFTPTNTISV 1320
 DB 1261 QFRLQGLGLOPFEGNVYFLEKRNQSEBVAVKPPLAENRNASIRDTFTPTNTISV 1320
 QY 1321 KPYFVNVYDVSNANVOTTVNLTVLQOPFGYWKQKAVGLKAEILHFOISAFISKSGSGL 1380
 DB 1321 KPYFVNVYDVSNANVOTTVNLTVLQOPFGYWKQKAVGLKAEILHFOISAFISKSGSGL 1380
 QY 1381 GRCQNVGVKLGYSR 1394
 DB 1381 GRCQNVGVKLGYSR 1394

RESULT 5

US-10-080-505-7

; Sequence 7, Application US/10080505

; Patent No. 6676948

GENERAL INFORMATION:
 APPLICANT: St. Gene, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/REF/DCF/DRR
 CURRENT APPLICATION NUMBER: US/10/080,505
 CURRENT FILING DATE: 2002-02-22
 PRIOR APPLICATION NUMBER: US 08/296,791
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 7
 LENGTH: 1395
 TYPE: PRT
 ORGANISM: Haemophilus influenzae
 US-10-080-505-7

Query Match 99.6%; Score 7246.5; DB 4; Length 1395;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1393; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 1 MKKTVFRLNPLTACTSIGIVSOAMAGTYGIDYOYRPFPAENKGFYGAOKITKYNQ 60
 DB 1 MKKTVFRLNPLTACTSIGIVSOAMAGTYGIDYOYRPFPAENKGFYGAOKITKYNQ 60
 QY 61 GOLVTSKTAPEMIDSVSRNGVAALYENQYIVSAHVNGTTVDVFGABGNPDQHRT 120
 DB 61 GOLVTSKTAPEMIDSVSRNGVAALYENQYIVSAHVNGTTVDVFGABGNPDQHRT 120
 QY 121 YKIVRNNYKKDMHVEDYDYNPLKFTYEAAPIDMTSNMGSTYSDTKYPERVRIG 180
 DB 121 YKIVRNNYKKDMHVEDYDYNPLKFTYEAAPIDMTSNMGSTYSDTKYPERVRIG 180
 QY 181 SCROFWRNDODKQVAGAYHTLZAGNTHNQAGANGYSYJAGVYRAGEYGEPLTASGK 240
 DB 181 SCROFWRNDODKQVAGAYHTLZAGNTHNQAGANGYSYJAGVYRAGEYGEPLTASGK 240
 QY 241 GSGSEWFTYDAKOKMLINGILRBGNPFEGKENGFOUVRKSYDEIFERDLHTSLYTRA 300
 DB 241 GSGSEWFTYDAKOKMLINGILRBGNPFEGKENGFOUVRKSYDEIFERDLHTSLYTRA 300
 QY 301 GNGVYTTISGNDNGQSLTQXSGISEIKITLANKSLPLKEXDKVHNRYDGPATYSPRLN 360
 DB 301 GNGVYTTISGNDNGQSLTQXSGISEIKITLANKSLPLKEXDKVHNRYDGPATYSPRLN 360
 QY 361 NGETLYTMDQKQSLIFASDINOAGGLYFEGNFTVSPNSQVTOQAGIHVSNSVTYWK 420
 DB 361 NGETLYTMDQKQSLIFASDINOAGGLYFEGNFTVSPNSQVTOQAGIHVSNSVTYWK 420
 QY 421 VNGVEHRLSKIGKGLTHVQAKGKNGKSIISVGDGVILLEQADDOQKQAFSEIGLVSGR 480
 DB 421 VNGVEHRLSKIGKGLTHVQAKGKNGKSIISVGDGVILLEQADDOQKQAFSEIGLVSGR 480
 QY 481 GTVQLNDCKQFDTDKFFGFRGRLDLNGHSLTFKRIQNTDEGAMTVNHTTOAANVTIT 540
 DB 481 GTVQLNDCKQFDTDKFFGFRGRLDLNGHSLTFKRIQNTDEGAMTVNHTTOAANVTIT 540
 QY 541 GNEISIVLPNGNNINKLDRKEIAVNGFETDKNHNRNLTYKPTEDRLLLSGNTN 600
 DB 541 GNEISIVLPNGNNINKLDRKEIAVNGFETDKNHNRNLTYKPTEDRLLLSGNTN 600
 QY 601 LKGDITQTKKLFPSGRPTPHAVNHLNKRSEMEGIPGGEIVWDHWINRTKAKENFOIK 660
 DB 601 LKGDITQTKKLFPSGRPTPHAVNHLNKRSEMEGIPGGEIVWDHWINRTKAKENFOIK 660
 QY 661 GGSAYVSRVSSIEGNMTVSNNAATFGVVPNOQNTICTSDMTGLTTCQKVDLDTKVI 720
 DB 661 GGSAYVSRVSSIEGNMTVSNNAATFGVVPNOQNTICTSDMTGLTTCQKVDLDTKVI 720
 QY 721 NSIPKTOINGSINTLDNATANVGLAKNGVNTLTNHSOFTLSNNAVOIGNIRLSDNSTA 780
 DB 721 NSIPKTOINGSINTLDNATANVGLAKNGVNTLTNHSOFTLSNNAVOIGNIRLSDNSTA 780

```

Qy 781 TYDNANLNGNVHLTDSAQSLKNSHSHQIGKGTVTLLNATWSPDTLLQNLN 840
Db 781 TYDNANLNGNVHLTDSAQSLKNSHSHQIGKGTVTLLNATWSPDTLLQNLN 840
Qy 841 STTLNSAYKSSNNTP--RRSLTETPTTSAHSPNTLVNGLSGQGTFOPTSSJFG 898
Db 841 STTLNSAYKSSNNTP--RRSLTETPTTSAHSPNTLVNGLSGQGTFOPTSSJFG 900
Qy 899 YKSDKLKLSNDACDYLTSVRNTGKEPTELBOJLTWESKDNQPSLSDKATLLNHDVAG 958
Db 901 YKSDKLKLSNDACDYLTSVRNTGKEPTELBOJLTWESKDNQPSLSDKATLLNHDVAG 960
Qy 959 ALRYLVZNDGEFRLNHPKKEQELNDLVBAEQAEKTLAKQVPTAKTGTGEPVRSRR 1018
Db 961 ALRYLVZNDGEFRLNHPKKEQELNDLVBAEQAEKTLAKQVPTAKTGTGEPVRSRR 1020
Qy 1019 AARAFPTLPPDQSLNLALEAKQAEKTLAKQVPTAKTGTGEPVRSRR 1078
Db 1021 AARAFPTLPPDQSLNLALEAKQAEKTLAKQVPTAKTGTGEPVRSRR 1080
Qy 1079 AALEYDAPQSSKEDLAQEAERKQKDLISRYNSALSSELATVNSMLSVQDELRL 1138
Db 1081 AALEYDAPQSSKEDLAQEAERKQKDLISRYNSALSSELATVNSMLSVQDELRL 1140
Qy 1139 EYDQASAVWNTAODKRRYSDAFRAYQOQXKTLROIQVOKALANGR-GAVFSHSDN 1198
Db 1141 EYDQASAVWNTAODKRRYSDAFRAY-QQKTLROIQVOKALANGR-GAVFSHSDN 1199
Qy 1199 TPDQVKNHATLTWMSGFAQVOQKDLQPGVNTGTSISAKRAEQRKIHRAINYVNA 1258
Db 1200 TPDQVKNHATLTWMSGFAQVOQKDLQPGVNTGTSISAKRAEQRKIHRAINYVNA 1259
Qy 1259 SYOPRLGQIGIOPFGVARYFERENYSEBVRKTPSLANVNGTIVYTPPTDNI 1318
Db 1260 SYOPRLGQIGIOPFGVARYFERENYSEBVRKTPSLANVNGTIVYTPPTDNI 1319
Qy 1319 SVKYEFFVNVVSNVQVTVNLVLOQPGFVWQKEVGLKALHFIQISAFISQGS 1378
Db 1320 SVKYEFFVNVVSNVQVTVNLVLOQPGFVWQKEVGLKALHFIQISAFISQGS 1379
Qy 1379 QLGKQNVGVKLGYSW 1394
Db 1380 QLGKQNVGVKLGYSW 1395

```

```

RESULT 6
US-10-080-505-9
: Sequence 9, Application US/10080505
: Patent No. 6576948
: GENERAL INFORMATION:
: APPLICANT: St. Geme, Joseph W.
: TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
: FILE REFERENCE: A-59941-1/Ref/DCP/DHR
: CURRENT APPLICATION NUMBER: US/10/080,505
: CURRENT FILING DATE: 2002-02-22
: PRIOR APPLICATION NUMBER: US 08/296,791
: PRIOR FILING DATE: 1994-10-25
: PRIOR APPLICATION NUMBER: US 09/839,996
: PRIOR FILING DATE: 2001-04-20
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 9
: LENGTH: 1434
: TYPE: PRT
: ORGANISM: Haemophilus influenzae
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (568)..(568)
: OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
: NAME/KEY: misc.feature
: LOCATION: (1702)..(1702)
: OTHER INFORMATION: "n" at position 1702 can be any base.

```

```

US-10-080-505-9
Query Match      84.1%; Score 6115; DB 4; Length 1434;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 13;

Qy 1 MKTVFRLNFLTACISIGIVSOAMAGHYEGIDVYRDEAKKGFVGAQNTKVNKQ 60
Db 1 MKTVFRLNFLTACISIGIVSOAMAGHYEGIDVYRDEAKKGFVGAQNTKVNKQ 60
Qy 61 GOLVTSMTKAMPIDFVSVRNGVAALVENVQYTVSVANTGTYDVDFGAEKNPQJREF 120
Db 61 GOLVTSMTKAMPIDFVSVRNGVAALVENVQYTVSVANTGTYDVDFGAEKNPQJREF 120
Qy 121 YKLVKNNYKKNLPEYEDYENHRLHKEVTEAPAPIMTSNNKGSYSRPTKPERPRIG 180
Db 121 YKLVKNNYKKNLPEYEDYENHRLHKEVTEAPAPIMTSNNKGSYSRPTKPERPRIG 180
Qy 181 SGQVFRANDQKQVAGAYHYITLANTHNGRAGANGSYSLGGDYAKAGEGPPPIAGSK 240
Db 181 SGQVFRANDQKQVAGAYHYITLANTHNGRAGANGSYSLGGDYAKAGEGPPPIAGSK 240
Qy 241 GDSGSPMTYDAEKQKWLINGILREBNPEKENGEPQVAKSYFDEIPERDLHLSITRA 300
Db 241 GDSGSPMTYDAEKQKWLINGILREBNPEKENGEPQVAKSYFDEIPERDLHLSITRA 300
Qy 301 GNGVYTTISGNDQGSITQKSGIPSEIKITLANNSLPLEKQKVENPRYDGPNIYSPRLN 360
Db 301 GNGVYTTISGNDQGSITQKSGIPSEIKITLANNSLPLEKQKVENPRYDGPNIYSPRLN 360
Qy 361 NSEFTLYFMDQKQSLIPASINQAGGILYFEGATTVSPNSNOTWQAGIHVSENSVYWK 420
Db 359 NSEFTLYFMDQKQSLIPASINQAGGILYFEGATTVSPNSNOTWQAGIHVSENSVYWK 418
Qy 421 VNGVEHRLSKIGLGLHVOAKGENKGSISYQKQKYLLEQCDQDQCNQQAASELGYLSR 480
Db 419 VNGVEHRLSKIGLGLHVOAKGENKGSISYQKQKYLLEQCDQDQCNQQAASELGYLSR 478
Qy 481 GTVQNDNDKQFDTKYFSGRGRDLNGLSHLTFKSIQNTDQAGMIVNHTTQAAVNTT 540
Db 479 GTVQNDNDKQFDTKYFSGRGRDLNGLSHLTFKSIQNTDQAGMIVNHTTQAAVNTT 538
Qy 541 GNESTVLP--NGNNTNKLDYRKELIANGMFGSETRKANGKRLNLYKPTTEDRLILSGGT 595
Db 533 GNESTVLP--NGNNTNKLDYRKELIANGMFGSETRKANGKRLNLYKPTTEDRLILSGGT 598
Qy 600 NLKQDITQTKGLTFSGRPTPHAVNHLNKRSMESGIPQGEIYMDHMIKRTKASNPDI 659
Db 599 NLKQDITQTKGLTFSGRPTPHAVNHLNKRSMESGIPQGEIYMDHMIKRTKASNPDI 658
Qy 660 KQGSAYSVRSVSSIEGNTVTSNNANATFGVVPNQDNTICTSDWTGLTTCQKVDLTDKV 719
Db 659 KQGSAYSVRSVSSIEGNTVTSNNANATFGVVPNQDNTICTSDWTGLTTCQKVDLTDKV 718
Qy 720 INSIPKQINGSLNLDAATANGKGLAKNGNTLVNHSOTFISNNATQICNIRLS---- 775
Db 719 INSIPKQINGSLNLDAATANGKGLAKNGNTLVNHSOTFISNNATQICNIRLS---- 778
Qy 776 ---DNST-----ATVDANLNGNVHLTDSAQSL 800
Db 779 ARVNNATLMDVNLADTSRFLTSNOATQIGTISLHQQAQATVDANLNGNVHLTDSAQSL 838
Qy 801 LKNSHSHQIGKGTVTLLNATWSPDTLLQNLNNTSTTLNSAYSASSNNTP--RR 859
Db 839 LKNSHSHQIGKGTVTLLNATWSPDTLLQNLNNTSTTLNSAYSASSNNTP--RR 858
Qy 860 RSLTETPTTSAHSPNTLVNGLSGQGTFOPTSSJFGYSKDLKLSNDACDYLTSVR 919
Db 899 RSLTETPTTSAHSPNTLVNGLSGQGTFOPTSSJFGYSKDLKLSNDACDYLTSVR 958
Qy 920 NTGKEPTELBOJLTWESKDNQPSLSDKATLLNHDVAGALRYLVZNDGEFRLNHPK 979
Db 959 NTGKEPTELBOJLTWESKDNQPSLSDKATLLNHDVAGALRYLVZNDGEFRLNHPK 1018

```

QY 98C QELNDLVAAEQAEERTLAKOYEPTACTGTGEBKYSRRARAAPPTLLPDOSLNTAE 1039
 Db 1019 QELSDLVRAEQAEERTLEAKYEQTAETQSNARVSR---RAVLSSTPSAQSLSNLALEV 1075
 QY 1040 KQALTAETQSKAKTKKYSRAV---FSDLLDQS-LFALAEALVLDADQSEKRL 1095
 Db 1076 KQABNAKTQSKAKTKKARSKRALREAFSDPTPLSQLNTLEALKVINAQPOKEKRO 1135
 QY 1096 AQ-SEAEKQKQKDLISRYNSALSLSATVSMLSYODEDLPLFDQOQSVMTNLAOD 1154
 Db 1136 AQSEAEKQKQKDLISRYNSALSLSATVSMLSYODEDLPLFDQOQSVMTNLAOD 1195
 QY 1155 KRRYSDAFRAYQOQKTLNARQIGYQALANGRIAGVFSHRSNDTDEQVKNHATLIMS 1214
 Db 1196 KRRYSDAFRAY-QQKTLNARQIGYQALANGRIAGVFSHRSNDTDEQVKNHATLIMS 1254
 QY 1215 GFAQYQMGDLQSGVNVGTGISASMAEQSRKIRKALINGVNASYQFPLQGLQPTLIG 1274
 Db 1255 GFAQYQMGDLQSGVNVGTGISASMAEQSRKIRKALINGVNASYQFPLQGLQPTLIG 1314
 QY 1275 VNRFFERENYQSEEVKQTPSLAFRRYNAGRVDYTPPTDNI SVKPYEFVNVDSNA 1334
 Db 1315 VNRFFERENYQSEEVKQTPSLAFRRYNAGRVDYTPPTDNI SVKPYEFVNVDSNA 1374
 QY 1335 NVQTYNLTVLQOQPFGRYQKEVGLKAEILHPIQSASFISQSGSLQKQNVGKGYRM 1394
 Db 1375 NVQTYNLTVLQOQSGRYQKEVGLKAEILHPIQSASFISQSGSLQKQNVGKGYRM 1434

RESULT 7

US-10-080-505-17
 ; Sequence 17, Application US/10080505
 ; Patent No. 6676948
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Geme, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/REF/DCF/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 17
 ; LENGTH: 1411
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; US-10-080-505-17

Query Match 86.1%; Score 5827.5; DB 4; Length 1411;
 Best Local Similarity 79.5%; Pred. No. 0;
 Matches 1133; Conservative 107; Mismatches 139; Indels 47; Gaps 12;

QY 1 MKKTYFRNLPLTACTSLGIVSOAMAGHYEGSIDYQYRRPEANRKGKTVGAQNTIKYKNO 60
 Db 1 MKKTYFRNLPLTACTSLGIVSOAMAGHYEGSIDYQYRRPEANRKGKTVGAQNTIKYKNO 60
 QY 61 GQLVSTMTKAPMIDSVSVSRNGVALVENVQYTVSAHVNVGYTVDFGAEGNPPDOHPT 120
 Db 61 GQLVSTMTKAPMIDSVSVSRNGVALVENVQYTVSAHVNVGYTVDFGAEGNPPDOHPT 120
 QY 121 YKIVRNNYKKDNHAYEDDYNPPLNHPYTFTEAPLMDTSMNGSTYSDTKYPERVIG 180
 Db 121 YKIVRNNYKKDNHAYEDDYNPPLNHPYTFTEAPLMDTSMNGSTYSDTKYPERVIG 180
 QY 181 SGRQFWNRNDQKEDQVAGAYHYLLTAGTNHQAGANGYSYGGVDRKAYGGLPLAGSK 240
 Db 181 SGRQFWNRNDQKEDQVAGAYHYLLTAGTNHQAGANGYSYGGVDRKAYGGLPLAGSK 240
 QY 241 GSGSGPMFYDAEKQKMLINGILBNGPFGKENGQFLVRKSYDEILFEEDLTLSTLYTRA 300

Db 241 GSGSGPMFYDAEKQKMLINGILBNGPFGKENGQFLVRKSYDEILFEEDLTLSTLYTRA 300
 QY 301 GNGVYTIISGNDGQSGITQSKGIPSEIKITLANNLPLEKQKYNHPYDQPNYSRPN 360
 Db 301 GNGVYTIISGNDGQSGITQSKGIPSEIKITLANNLPLEKQKYNHPYDQPNYSRPN 360
 QY 361 NGETLYFMDQKGSLLPASDINGAGSLYFEGNETVSPNSQOTWQAGIHVSENSIVTK 420
 Db 359 NGETLYFMDQKGSLLPASDINGAGSLYFEGNETVSPNSQOTWQAGIHVSENSIVTK 418
 QY 421 VNVCHEDRLSTIGGTLHYQAKENKSGSYQDQKVLLEQOADDQANKAFSEIGVSGR 480
 Db 419 VNVCHEDRLSTIGGTLHYQAKENKSGSYQDQKVLLEQOADDQANKAFSEIGVSGR 478
 QY 481 GTVQLNDKQPTDQPFPGRGRLDINGSLTEPKRIQNTDGAATVNHNTQAAKTTT 540
 Db 479 GTVQLNDKQPTDQPFPGRGRLDINGSLTEPKRIQNTDGAATVNHNTQAAKTTT 538
 QY 541 GNESTVLNPGANNINKLDYRELAIVNMFGEYTDKKNHGRLLTYPTTEDERTLISGGIN 600
 Db 539 GSDTL---NDNTGDLINKRDLAFNGMGQDQDQKQNTGRLANTVYPLKQKHPLISGIN 594
 QY 601 LKGDITQKRGKLPFSGRPTPHAYNHLAKGSEMEGIPQSEIYWDHM-INRTFAENPQIK 660
 Db 595 LKGNITQDQGLTVFSGRPTPHAYNHL-INTEL-GPKXEVVIDDM-INRTFAENPQIK 652
 QY 661 GGSAAVSNVNSIEGMMTVSNANAATGVVNOONTICRSPTWGLTTCQKQVLTQTKYI 720
 Db 653 GGSAAVSNVNSIEGMMTVSNANAATGVVNOONTICRSPTWGLTTCQKQVLTQTKYI 712
 QY 721 NSIPETQINSINLTUNATANYKGLAKNGVTLINHSQPTLSNNAQIGNIRLSNSTA 780
 Db 713 NSIPETQINSINLTUNATANYKGLAKNGVTLINHSQPTLSNNAQIGNIRLSNSTA 772
 QY 781 TYDANANGVHLTDSAPSLKNSHFSHQIQKKTGTYTLENATYTPSTTLQNTLTK 840
 Db 773 TYDANANGVHLTDSAPSLKNSHFSHQIQKKTGTYTLENATYTPSTTLQNTLTK 832
 QY 842 STTLNSAYASSNNTPR-RSLSETTTPSAHRPNTLYNGKLSQGFQGTSLSEY 899
 Db 833 STTLNSAYASSNNTPR-RSLSETTTPSAHRPNTLYNGKLSQGFQGTSLSEY 892
 QY 900 KSDKIKLSNABCDYILSVRNTGKEPTELOLTVESKONOPLDKLTLENDHVDGA 959
 Db 893 KSDKIKLSNABCDYILSVRNTGKEPTELOLTVESKONOPLDKLTLENDHVDGA 952
 QY 960 LRYKLVNDGEFPLNPIKROELNDLVAAEQAEERTLAKOYEPTACTGTGEBKYSRRA 1019
 Db 953 LRYKLVNDGEFPLNPIKROELNDLVAAEQAEERTLAKOYEPTACTGTGEBKYSRRA 1010
 QY 1020 ARAAPPTLLPDOSLNALEAKQAELETAETQSKA-----KTKYVSRK---AV 1064
 Db 1011 ARAAPPTLLPDOSLNALEAKQAELETAETQSKA-----KTKYVSRK---AV 1069
 QY 1065 FSDLLDQS---FA-----LEALAEVLDADQSEKRLAQSEAEKQKQKD 1108
 Db 1070 FSDLLDQS---FA-----LEALAEVLDADQSEKRLAQSEAEKQKQKD 1126
 QY 1109 LISRYNSALSLSATVSMLSYODEDLPLFDQOQSVMTNLAODKRRYSDAFRAYQ 1168
 Db 1127 LISRYNSALSLSATVSMLSYODEDLPLFDQOQSVMTNLAODKRRYSDAFRAYQ 1185
 QY 1169 QKTLNARQIGYQALANGRIAGVFSHRSNDTDEQVKNHATLIMSAGFQYQMGDLQPGV 1228
 Db 1186 QKTLNARQIGYQALANGRIAGVFSHRSNDTDEQVKNHATLIMSAGFQYQMGDLQPGV 1245
 QY 1229 NVGNGISASMAEQSRKIRKALINGVNASYQFPLQGLQPTLIGSVNRYFFERENYQSE 1288
 Db 1246 NVGNGISASMAEQSRKIRKALINGVNASYQFPLQGLQPTLIGSVNRYFFERENYQSE 1305
 QY 1289 EVRYKTPSLAFRRYNAGRVDYTPPTDNI SVKPYEFVNVDSNANVQTYNLTLYQOP 1348

Db 1306 EYKVCPTSLAFNRNAGVRVDYTFPTENISVKEFFVNVYDVSNANQTVRAVLQGP 1365

QY 1349 FGRYWCXKVEGKAEILHFOISAFISKSQSLQKQONVGVKXGYW 1394

Db 1366 FGRYWCXKVEGKAEILHFOISAFISKSQSLQKQONVGVKXGYW 1411

RESULT 8

US-10-080-505-13

/ Sequence 13, Application US/10080505

/ Patent No. 6676948

/ GENERAL INFORMATION:

/ APPLICANT: St. Gene, Joseph W.

/ TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS

/ FILE REFERENCE: A-59941-1/PT/DCF/CHR

/ CURRENT APPLICATION NUMBER: US/10/080,505

/ CURRENT FILING DATE: 2002-02-22

/ PRIORITY FILING DATE: US 08/296,791

/ PRIORITY FILING DATE: 1994-10-25

/ PRIOR APPLICATION NUMBER: US 09/839,996

/ PRIOR FILING DATE: 2001-04-20

/ NUMBER OF SEQ ID NOS: 58

/ SOFTWARE: Patent version 3.1

/ SEQ ID NO 13

/ LENGTH: 1436

/ TYPE: PRT

/ ORGANISM: Haemophilus influenzae

US-10-080-505-13

Query Match 79.8%; Score 5807; DB 4; Length 1436;

Best Local Similarity 79.3%; Pred. No. 0;

Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;

QY 1 MKKTVFLNLTACISLGIVSOAAGHTYGDYQYRDPFENKGFYTGACNTKYKQ 60

Db 1 MKKTVFLNLTACISLGIVSOAAGHTYGDYQYRDPFENKGFYTGACNTKYKQ 60

QY 61 GQVYGSMTKAPMIDSVYSRNGVALVENQYIVSAHVAYTDVDFGAEKGNPQHRFT 120

Db 61 GQVYGSMTKAPMIDSVYSRNGVALVENQYIVSAHVAYTDVDFGAEKGNPQHRFT 120

QY 121 YKIVKNNYKCNLHPEYEDDYNHRLKPYTAALITMTSNNGSYSDRTYPERVATG 180

Db 121 YKIVKNNYKCNLHPEYEDDYNHRLKPYTAALITMTSNNGSYSDRTYPERVATG 180

QY 121 YQIVKNNYKCNLHPEYEDDYNHRLKPYTAALITMTSNNGSYSDRTYPERVATG 180

Db 121 YQIVKNNYKCNLHPEYEDDYNHRLKPYTAALITMTSNNGSYSDRTYPERVATG 180

QY 181 SGQRFENDD-----KGDYVAGAYHLLTAKGTHNORAGAGSYSGVGVKRAEYEPD 235

Db 181 SGQRFENDD-----KGDYVAGAYHLLTAKGTHNORAGAGSYSGVGVKRAEYEPD 235

QY 236 IAGSKDSSPMFIVDAEKOKMLINGILBEGNPEKANGFQIVRKSYF-DEIFERDLHT 294

Db 236 IAGSKDSSPMFIVDAEKOKMLINGILBEGNPEKANGFQIVRKSYF-DEIFERDLHT 294

QY 241 TGVSFGDSGPMFIVDAIKQKWLINGVLQTNPFSGANGFQILRNKMFYDNYVEDLPI 300

Db 241 TGVSFGDSGPMFIVDAIKQKWLINGVLQTNPFSGANGFQILRNKMFYDNYVEDLPI 300

QY 295 SLVFRANGVYVITISNDNGQSGITQ--KSGIPSEIKITLANNSLPLEKOKVHNPRY-- 349

Db 295 SLVFRANGVYVITISNDNGQSGITQ--KSGIPSEIKITLANNSLPLEKOKVHNPRY-- 349

QY 301 TYLEPNSKGHISFTISNNNGTGTVQTNKCVSPQFKVYRTQVLNENALKEKDK--EPYAA 358

Db 301 TYLEPNSKGHISFTISNNNGTGTVQTNKCVSPQFKVYRTQVLNENALKEKDK--EPYAA 358

QY 350 DGPNIYSPRIANGELIYFMQKQSLIFASDINQAGSLFEGNPFVSPNSNQWQAGI 405

Db 350 DGPNIYSPRIANGELIYFMQKQSLIFASDINQAGSLFEGNPFVSPNSNQWQAGI 405

QY 359 GGVNAYKPRILANGKXIYFGDRGTGLITINNINQAGSLFEGNPFVSPNSNNAWQAGAG 413

Db 359 GGVNAYKPRILANGKXIYFGDRGTGLITINNINQAGSLFEGNPFVSPNSNNAWQAGAG 413

QY 410 HVSENGSTVWKVNGVEHDLSTKIGKTLHVAQKENGKSIWAGKVLLEQADDOGNKQ 469

Db 410 HVSENGSTVWKVNGVEHDLSTKIGKTLHVAQKENGKSIWAGKVLLEQADDOGNKQ 469

QY 419 HVSENGSTVWKVNGVEHDLSTKIGKTLHVAQKENGKSIWAGKVLLEQADDOGNKQ 478

Db 419 HVSENGSTVWKVNGVEHDLSTKIGKTLHVAQKENGKSIWAGKVLLEQADDOGNKQ 478

QY 470 AFSEGLVSGRGTVQLNDNDQEDTDEKYGGRGRGLNLNHSITTKRIQNTDGAIVNH 529

Db 470 AFSEGLVSGRGTVQLNDNDQEDTDEKYGGRGRGLNLNHSITTKRIQNTDGAIVNH 529

QY 479 AFKEVGIYSGRATVQLNSADQVDPNNIYFGRGRGLDLNHSITTKRIQNTDGAIVNH 538

Db 479 AFKEVGIYSGRATVQLNSADQVDPNNIYFGRGRGLDLNHSITTKRIQNTDGAIVNH 538

QY 530 NTTQANVITIGNESVLP--NGNNIKLIDYKKEJAYVNGFGEEDKKNKNGNLITKPTT 588

Db 530 NTTQANVITIGNESVLP--NGNNIKLIDYKKEJAYVNGFGEEDKKNKNGNLITKPTT 588

QY 539 NTTQANVITIGNESVLP--NGNNIKLIDYKKEJAYVNGFGEEDKKNKNGNLITKPTT 598

Db 539 NTTQANVITIGNESVLP--NGNNIKLIDYKKEJAYVNGFGEEDKKNKNGNLITKPTT 598

QY 589 EDRLTLISGCTNKQDITQTKGKLFPSGRPTPAVNHAKWSEMGIDQGEIWDHDM 648

Db 589 EDRLTLISGCTNKQDITQTKGKLFPSGRPTPAVNHAKWSEMGIDQGEIWDHDM 648

QY 599 EDRLTLISGCTNKQDITQTKGKLFPSGRPTPAVNHAKWSEMGIDQGEIWDHDM 656

Db 599 EDRLTLISGCTNKQDITQTKGKLFPSGRPTPAVNHAKWSEMGIDQGEIWDHDM 656

QY 649 NRTKAEIPQIGGSANVSRVSSIEKWTVSNNANATFGVVPQONNTICTSDWTGLTT 708

Db 649 NRTKAEIPQIGGSANVSRVSSIEKWTVSNNANATFGVVPQONNTICTSDWTGLTT 708

QY 657 TRTKAEIPQIGGSANVSRVSSIEKWTVSNNANATFGVVPQONNTICTSDWTGLTT 716

Db 657 TRTKAEIPQIGGSANVSRVSSIEKWTVSNNANATFGVVPQONNTICTSDWTGLTT 716

QY 709 CQKVDLIDTVEVINSIFPTQINGSINTDNATANYGAKNGVTLTNNISQPTSSNNA 768

Db 709 CQKVDLIDTVEVINSIFPTQINGSINTDNATANYGAKNGVTLTNNISQPTSSNNA 768

QY 717 CXTVDLIDTVEVINSIFPTQINGSINTDNATANYGAKNGVTLTNNISQPTSSNNA 776

Db 717 CXTVDLIDTVEVINSIFPTQINGSINTDNATANYGAKNGVTLTNNISQPTSSNNA 776

QY 769 IGTIELSDNSTATVDNANANGNHLTDSQFSLKNSHSHQIQDGTIVTLNNAITMP 828

Db 769 IGTIELSDNSTATVDNANANGNHLTDSQFSLKNSHSHQIQDGTIVTLNNAITMP 828

QY 777 TGNIOJLSNANATVDNANANGNHLTDSQFSLKNSHSHQIQDGTIVTLNNAITMP 836

Db 777 TGNIOJLSNANATVDNANANGNHLTDSQFSLKNSHSHQIQDGTIVTLNNAITMP 836

QY 829 SDTLNNTLANSSTTLNNSAYASANNTR--RSLFETTTPTSAHRFVTLTVNGKLSG 887

Db 829 SDTLNNTLANSSTTLNNSAYASANNTR--RSLFETTTPTSAHRFVTLTVNGKLSG 887

QY 837 SDTLNNTLANSSTTLNNSAYASANNTR--RSLFETTTPTSAHRFVTLTVNGKLSG 896

Db 837 SDTLNNTLANSSTTLNNSAYASANNTR--RSLFETTTPTSAHRFVTLTVNGKLSG 896

QY 888 GTFOPTSSLFYKSDKXKLSNDAEGDYILSVANNKEPETLEQTLTVESKONOP--SDKX 947

Db 888 GTFOPTSSLFYKSDKXKLSNDAEGDYILSVANNKEPETLEQTLTVESKONOP--SDKX 947

QY 897 GTFOPTSSLFYKSDKXKLSNDAEGDYILSVANNKEPETLEQTLTVESKONOP--SDKX 956

Db 897 GTFOPTSSLFYKSDKXKLSNDAEGDYILSVANNKEPETLEQTLTVESKONOP--SDKX 956

QY 948 FTLENDHNDAGALKYKXNDGFEFLNPIYQELHNDIVRABQARTLEAKOVEPTAKT 1007

Db 948 FTLENDHNDAGALKYKXNDGFEFLNPIYQELHNDIVRABQARTLEAKOVEPTAKT 1007

QY 957 FTLENDHNDAGALKYKXNDGFEFLNPIYQELHNDIVRABQARTLEAKOVEPTAKT 1016

Db 957 FTLENDHNDAGALKYKXNDGFEFLNPIYQELHNDIVRABQARTLEAKOVEPTAKT 1016

QY 1008 QTGEPTKRSRABAPDPTLPDQSLNANLAKAKHETLEQTSKAKTKKYSKRAVFS 1067

Db 1008 QTGEPTKRSRABAPDPTLPDQSLNANLAKAKHETLEQTSKAKTKKYSKRAVFS 1067

QY 1017 QTNAKAVRK--RAVSDTLTPDQSLDVLQAEQVEPTAKKONNA--KKRSRAVFS 1071

Db 1017 QTNAKAVRK--RAVSDTLTPDQSLDVLQAEQVEPTAKKONNA--KKRSRAVFS 1071

QY 1068 PLIDQS-----LEALEALEVID 1085

Db 1068 PLIDQS-----LEALEALEVID 1085

QY 1072 TLPDQSLDVLQAEQVEPTAKKONNAKAKKAKKAAEFSDPTLQSLKYLEVLRV 1131

Db 1072 TLPDQSLDVLQAEQVEPTAKKONNAKAKKAKKAAEFSDPTLQSLKYLEVLRV 1131

QY 1086 APQOSKRLQAEBAEKOKDLISRYNSALSSELSATVNSLSVQDELDTLPVDAQS 1145

Db 1086 APQOSKRLQAEBAEKOKDLISRYNSALSSELSATVNSLSVQDELDTLPVDAQS 1145

QY 1132 AQQVQKKE--PDQEKOKKDLISRYNSALSSELSATVNSLSVQDELDTLPVDAQS 1188

Db 1132 AQQVQKKE--PDQEKOKKDLISRYNSALSSELSATVNSLSVQDELDTLPVDAQS 1188

QY 1146 AWTNIAQDKRYSYDPAFRAYQOQKTNRLQIGVQYKALANGRGVFSHSBDMTPBOYK 1205

Db 1146 AWTNIAQDKRYSYDPAFRAYQOQKTNRLQIGVQYKALANGRGVFSHSBDMTPBOYK 1205

QY 1189 AWTNIAQDKRYSYDPAFRAYQOQKTNRLQIGVQYKALANGRGVFSHSBDMTPBOYK 1247

Db 1189 AWTNIAQDKRYSYDPAFRAYQOQKTNRLQIGVQYKALANGRGVFSHSBDMTPBOYK 1247

QY 1206 NHALTLMNSGPAQYQWGDQFGVNVGTGISASMAEBSQKILHRAKLVGNASVYFRLG 1265

Db 1206 NHALTLMNSGPAQYQWGDQFGVNVGTGISASMAEBSQKILHRAKLVGNASVYFRLG 1265

QY 1248 NHALTLMNSGPAQYQWGDQFGVNVGTGISASMAEBSQKILHRAKLVGNASVYFRLG 1307

Db 1248 NHALTLMNSGPAQYQWGDQFGVNVGTGISASMAEBSQKILHRAKLVGNASVYFRLG 1307

QY 1266 QLGIGPYGVNRYFIERENYQSEBYRYKTSPLAFNRYNAGLRVDYTFPTDNI SVAPYF 1335

Db 1266 QLGIGPYGVNRYFIERENYQSEBYRYKTSPLAFNRYNAGLRVDYTFPTDNI SVAPYF 1335

QY 1308 QLGIGPYGVNRYFIERENYQSEBYRYKTSPLAFNRYNAGLRVDYTFPTDNI SVAPYF 1367

Db 1308 QLGIGPYGVNRYFIERENYQSEBYRYKTSPLAFNRYNAGLRVDYTFPTDNI SVAPYF 1367

QY 1326 VNVYDVSNANQTVRAVLQGPGRYWCXKVEGKAEILHFOISAFISKSQSLQKQON 1385

Db 1326 VNVYDVSNANQTVRAVLQGPGRYWCXKVEGKAEILHFOISAFISKSQSLQKQON 1385

QY 1368 VNVYDVSNANQTVRAVLQGPGRYWCXKVEGKAEILHFOISAFISKSQSLQKQON 1427

Db 1368 VNVYDVSNANQTVRAVLQGPGRYWCXKVEGKAEILHFOISAFISKSQSLQKQON 1427

QY 1386 VGVKXGYW 1394

Db 1386 VGVKXGYW 1436

RESULT 9

US-10-080-505-15

/ Sequence 15, Application US/10080505

/ Patent No. 6676948

/ GENERAL INFORMATION:

/ APPLICANT: St. Gene, Joseph W.

/ TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS

/ FILE REFERENCE: A-59941-2/PT/DCF/CHR

/ CURRENT APPLICATION NUMBER: US/10/080,505

/ PRIOR FILING DATE: 2002-02-22

/ PRIOR APPLICATION NUMBER: US 08/296,791

; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 1391
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; US-10-080-505-15

Query Match 72.9%; Score 5303.5; DB 4; Length 1391;
 Best Local Similarity 74.8%; Pred. No. 0;
 Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

QY 1 MKTVRLNPLTACISLIGIVSQAAMAGHYEGIDYQYRDAENKGFYVGAQNIKYNNKQ 60
 DB 1 MKTVRLNPLTACISLIGIVSQAAMAGHYEGIDYQYRDAENKGFYVGAQNIKYNNK 60
 QY 61 GQLVGSKTKAPMIDFSVVSNGVAALVENQYIVSAHVNGYTDVDFGAEGNPDQHRFT 120
 DB 61 GEMIGTMMGVPMPLDSMVRGGYSTLSSEQHLISVAHVNGYTDVDFGEGENPDQHRFX 120
 QY 121 YKIVKNNYKKNLHPEFDYHNPRLHFKVTBAAPIDMTSNXGSSYSDRTKYPERVRIQ 180
 DB 121 YKIVKNNYKKNLHPEFDYHNPRLHFKVTBAAPIDMTSNXGSSYSDRTKYPERVRIQ 180
 QY 121 YKIVKNNYKKNLHPEFDYHNPRLHFKVTBAAPIDMTSNXGSSYSDRTKYPERVRIQ 180
 DB 121 YKIVKNNYKKNLHPEFDYHNPRLHFKVTBAAPIDMTSNXGSSYSDRTKYPERVRIQ 180
 QY 181 SGHOFNRNDQK--GDQVAGAYATLTAGTINQAGAGNGSYSLGQDVKAGEVPLPIAG 238
 DB 179 SGHOFNRNDQK--GDQVAGAYATLTAGTINQAGAGNGSYSLGQDVKAGEVPLPIAG 238
 QY 239 SKGDSSEPFYIDAEKOKMLINGILREGNPEEGENGFOIVKSYFDEIFERDLFTSLYT 298
 DB 238 SKGDSSEPFYIDAEKOKMLINGILREGNPEEGENGFOIVKSYFDEIFERDLFTSLYT 298
 QY 238 SFQDSSEPFYIDAEKOKMLINGILREGNPEEGENGFOIVKSYFDEIFERDLFTSLYT 298
 DB 238 SFQDSSEPFYIDAEKOKMLINGILREGNPEEGENGFOIVKSYFDEIFERDLFTSLYT 298
 QY 299 BAANGYTTISANDG-----QGSITQKSGIPSEIKITLANGL--PLKEDKVNPRIDGP 352
 DB 298 BAANGYTTISANDG-----QGSITQKSGIPSEIKITLANGL--PLKEDKVNPRIDGP 352
 QY 353 NITYSPRLNNGETLYFMOQKQSLIFASDINQAGAGLYFGSNTTYSFNSQ--TNGQGHY 411
 DB 358 NITYSPRLNNGETLYFMOQKQSLIFASDINQAGAGLYFGSNTTYSFNSQ--TNGQGHY 411
 QY 412 SENSIVTWKNGVHDLRLKIQKGLTAVQAKENGKSIYVQKGLILEEQADQKQKQAF 471
 DB 418 SENSIVTWKNGVHDLRLKIQKGLTAVQAKENGKSIYVQKGLILEEQADQKQKQAF 471
 QY 472 SEIGVSGRGVQVQNDKQPDYDKYFPGFGRG--LJNGSLTFKKIQTDBGAMIVNHT 531
 DB 478 SEIGVSGRGVQVQNDKQPDYDKYFPGFGRG--LJNGSLTFKKIQTDBGAMIVNHT 531
 QY 532 TQAAVTTIGNESIVLPNGNINIKLKYKELIAYVNGFETDKNKNRRLMLIYKPTEDR 591
 DB 538 TQAAVTTIGNESIVLPNGNINIKLKYKELIAYVNGFETDKNKNRRLMLIYKPTEDR 591
 QY 592 TLISGGTNLKGDTYQTKRPFSGRPPTPAVNHNLKEMSEWEGIPQELIWDHMIKRT 651
 DB 594 TLISGGTNLKGDTYQTKRPFSGRPPTPAVNHNLKEMSEWEGIPQELIWDHMIKRT 651
 QY 652 FKAENFOIKGSSAVSRNYSIEGHWYVSNNAAITGVVPCOQNTICTRSDDMTGLTQK 711
 DB 654 FKAENFOIKGSSAVSRNYSIEGHWYVSNNAAITGVVPCOQNTICTRSDDMTGLTQK 711
 QY 712 VDLDTKYNSIPKTOINGISILDNATAVNGGLATLNGNTLVNHSOFTLSNNTQIGN 771
 DB 714 VDLDTKYNSIPKTOINGISILDNATAVNGGLATLNGNTLVNHSOFTLSNNTQIGN 771
 QY 772 ILSNSTAVVDNANNGVNLTPSAQFSLSNHSFHOQDQKGTVTLLENATWTPSDY 831
 DB 774 ILSNSTAVVDNANNGVNLTPSAQFSLSNHSFHOQDQKGTVTLLENATWTPSDY 831
 QY 832 TLQNTLANSSTILNAAYSASNNTP--RRSLFTTTTTSAAHRENTLTVNGKISGGGT 889
 DB 834 TLQNTLANSSTILNAAYSASNNTP--RRSLFTTTTTSAAHRENTLTVNGKISGGGT 889

QY 890 FQTSLSFGYKSDKLKISNDAGDYILSVRNTGKEPTLEQUTLVESKONQPLSDKLXT 949
 DB 894 FQTSLSFGYKSDKLKISNDAGDYILSVRNTGKEPTLEQUTLVESKONQPLSDKLXT 949
 QY 950 LENDHYDAGALRYKLVNNGEERLNPKEQELHNDLVNABAEKTELEAKQVETAKTOT 1009
 DB 954 LENDHYDAGALRYKLVNNGEERLNPKEQELHNDLVNABAEKTELEAKQVETAKTOT 1009
 QY 1010 GEPKYSRRARAAPETLPDQSLNALAEKQELTAETQSKATKYRSKRAV--FSD 1067
 DB 1014 GEPKYSRRARAAPETLPDQSLNALAEKQELTAETQSKATKYRSKRAV--FSD 1067
 QY 1068 PLDQSLFALAEVTDAPQSEKORLQAEAK--QKQKDLISRYNSNLSLSTATVN 1126
 DB 1068 PLDQSLFALAEVTDAPQSEKORLQAEAK--QKQKDLISRYNSNLSLSTATVN 1126
 QY 1127 SMLSYDELDRLPVDAQSAVWNTNACQKRYSDAFRAYQOQKTNLROIGVQKALANR 1186
 DB 1125 SMLSYDELDRLPVDAQSAVWNTNACQKRYSDAFRAYQOQKTNLROIGVQKALANR 1186
 QY 1187 IGAVERSHSRSDMTEDQVQKHAATLTNMSGFAQYQWGLQFGVNVGTGSIASKAAEQSRK 1246
 DB 1184 IGAVERSHSRSDMTEDQVQKHAATLTNMSGFAQYQWGLQFGVNVGTGSIASKAAEQSRK 1246
 QY 1247 IIRKAINYNASVQYRLQGLQIGPYGVNRYFERENTQSEEVRYVTPSLAFNRNAGI 1306
 DB 1244 IIRKAINYNASVQYRLQGLQIGPYGVNRYFERENTQSEEVRYVTPSLAFNRNAGI 1306
 QY 1307 RYDYTFPTNINISVRYKFFPVNVVNSANVQTVNLTVIQQEPRYQKQVGAAILHFE 1366
 DB 1304 RYDYTFPTNINISVRYKFFPVNVVNSANVQTVNLTVIQQEPRYQKQVGAAILHFE 1366
 QY 1367 QLSAFISKSQSQSLGKQNVGVKLGYNW 1394
 DB 1364 QLSAFISKSQSQSLGKQNVGVKLGYNW 1394

RESULT 10
 US-10-080-505-11
 ; Sequence 11, Application US/10080505
 ; Patent No. 6676948
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO 11
 ; LENGTH: 1391
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; US-10-080-505-11

Query Match 72.8%; Score 5297.5; DB 4; Length 1391;
 Best Local Similarity 74.7%; Pred. No. 0;
 Matches 1052; Conservative 109; Mismatches 216; Indels 31; Gaps 15;

QY 1 MKTVRLNPLTACISLIGIVSQAAMAGHYEGIDYQYRDAENKGFYVGAQNIKYNNKQ 60
 DB 1 MKTVRLNPLTACISLIGIVSQAAMAGHYEGIDYQYRDAENKGFYVGAQNIKYNNK 60
 QY 61 GQLVGSKTKAPMIDFSVVSNGVAALVENQYIVSAHVNGYTDVDFGAEGNPDQHRFT 120
 DB 61 GEMIGTMMGVPMPLDSMVRGGYSTLSSEQHLISVAHVNGYTDVDFGEGENPDQHRFX 120
 QY 121 YKIVKNNYKKNLHPEFDYHNPRLHFKVTBAAPIDMTSNXGSSYSDRTKYPERVRIQ 180

```

Db      1184 IGAVESHSSDNTDEQVKNHATLMMSGRQGWMDIOPVWVGAGIASMAMEDSKR 1245
Qy      1241 IHRKALNYGNASQVPLGQLGIDYFGVNRVFIKENYQSEBRYKTPSLAFNRNAGI 1306
Db      1244 IHRKALNYGNASQVPLGQLGIDYFGVNRVFIKENYQSEBRYKTPSLAFNRNAGI 1303
Qy      1307 RVDTFPEPNDISVWPEFENRVVDVSNANQVTVNLTLVQPFERRWQGEVGLAASLTHF 1366
Db      1304 RVDTFPEPNDISVWPEFENRVVDVSNANQVTVNRTMLQOSFERWQGEVGLAASLTHF 1363
Qy      1367 QISAFISKSGSQSGQKQONVYGLCYRM 1394
Db      1364 QISAFISKSGSQSGQKQONVYGLCYRM 1391

RESULT 11
US-08-296-791-3
; Sequence 3, Application US/08296791
; Patent No. 6245337
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tregearth, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1541 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; US-08-296-791-3

Query March 23.0%; Score 1673.5; DB 3; Length 1541;
Best Local Similarity 28.5%; Pred. No. 1.4e-101;
Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

Qy      6 FRNLPTACISLIGVQAMAGHTFGSDYQYRDPAENKGFVVGAKHLYXWKKQGLTG 65
Db      6 FKNLPTALTYALVYTTTAAVADVDVQIFRPAENKSKSVGATVTVLVKDKNKVDIG 65
Qy      66 TSWTGA-PMIDFSYVS-RNGVALVENVQTVSVAH-NVGYTDVDFG-AEGN----NPDH 117
Db      66 TALPGLIMIDIFSVDVDRKATLILNPQYVVGKHSNGVSEILHGKLNQNMNNNAKXAH 125
Qy      118 RFT-----YKIVKNNY-----KKDLNHYEDDVHNPRLHKFTVTEAAPID-MTS 160
Db      126 RVDSSENRKRYSVEKQRIYTKLNKRYTTFEDQYQRRREDYTPRLDKFTVLEVAPISASTA 185
Qy      161 NMNNGSTYSDFRTKPERVAVIGSGRQFWRNDQDKDQ-----VAGAVHVL 203

```


SEQUENCE DESCRIPTION: SEQ ID NO: 3
US-09-839-936-3

Query Match 23.3%; Score 1673.5; DB 4; Length 1541;
Best Local Similarity 28.5%; Pred. No. 1,4e-101;
Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

QY 6 PRLNLTACISLIGVSOAMAGHTFYGIDYQYDPAENKKGKFTVGAQNIKYNNKQGLVG 65
DB 6 PRLNLTACISLIGVSOAMAGHTFYGIDYQYDPAENKKGKFTVGAQNIKYNNKQGLVG 65
QY 66 TSMYKA-PMIDESVYS-RNGVAAVENVQIVSYAH-NVGYTDVDFG-ABGN---KPDQH 117
DB 66 TALNGLPMIDFSVVDVDRIALTLINPQYVGVKESNGVELFGLNGNNNNNNNAKAH 125
QY 118 RFT-----YKIVKANNY-----KXONIHPYEDDYNPNLHKFTYEAAPD-MTS 166
DB 126 RDVSSSENRKFSVEKKEPPTKLNKGTVTTEDOTQRREDYVPRDKFVTEVAPLEASTA 185
QY 161 NMNGSYSDRTKYPVVAIGSGRQFMHNDODKQD-----VAGAYHYL 203
DB 186 SSDGTYNDQNKYPAFVILSGSQF---LYKGDVYSLILNHEVGANNLKLVDGATTYG 242
QY 204 TACGTENQKAGKGYSTYG-----GVYKAGEYGLPLIAGSKDGSMPITYAKQKX 257
DB 243 IAGPYKVNHNNGILGFGNSKEHSDPKGILSODELTNYAVLGDGSPFLFYDEREKKA 302
QY 258 LINGILREGNPEFEKKE-NGFQYVKSYPDEIFERDLHSL-----YTAQNG-VTITSG 309
DB 303 LFGSDYDFMAGYKKSQWEMNITYKSOFTKQVANKQASGLSKTDYSSNSGKTSTTIG 362
QY 310 NDNGQSGITQKSGIPSEIKITLANMSLPLKEDKQVNPFRYDGNITVSPILNGETLYFM 369
DB 363 GE-----KSLNDVLAD-----GDKX-----NHGSYTF-- 386
QY 370 QKQSGILFASDINGQAGLYPEGNFTVSPNS-NQWQAGIHYVSNSTYVAKVNGVSHR 428
DB 387 EGSGTLTNNIDQAGGLPPEGDYEVKGTSDNITWKAGAVYABEKTYVWKVNPQYR 446
QY 429 LSKIGKGLTAVQAKENGKSGISYDGKVLLEQQLDQGNQASHLGLVSRGVQLAND 486
DB 447 LAKIGKGLTIVGTDQNGKSLKVGDDGTVLKQQTNGSG-QHAFASGIVSRSRTLYLAND 505
QY 489 KQPDYKXYFGRGRGLDNGSLPLFKRLQNTDEGAMIVNENYTOQAVNTITGNS-VLP 548
DB 506 KQYDPNSIYFGRGRGLDNGSLPLFKRLQNTDEGAMIVNENYTOQAVNTITGNS-VLP 548
QY 549 N----- 549
DB 566 NTITPYNIDAPEDNYPAFRIIDQGLYLMLENVYTYALRKQASTRELPKNSGESENK 625
QY 550 -----GNNIKLDYRKELIAYNGMFGETDKKNGRLNLYPKTTERKTLILSG 597
DB 626 WLYMKTSDPAKAVNYNHNNHBMNGFNGYCFE-BEKKNNGNLNVTFKSGSEONRRLILG 684
QY 598 GTYKAGDITQYKALFFSGPTPHAYNHLNRKWEKGI-----PGSEIYVDDHM 647
DB 685 GTYKAGDITQYKALFFSGPTPHAYNHLNRKWEKGI-----PGSEIYVDDHM 647
QY 648 INRTFAENFOIKGSGAVYS-RWSSISGKMTVYNNANNTFGVYVNOQNTITCRSPTGL 706
DB 737 INRTFAENFOIKGSGAVYS-RWSSISGKMTVYNNANNTFGVYVNOQNTITCRSPTGL 706
QY 707 TCGQKVDLTFRKINSIPKTOINGSLNLTJNATANYGLAKXNGVNTILHSGQFTLSNNA 766
DB 795 VITCTDKLSD-KALNSPNTNLRGNVLTRESANF-VLGAANIPLGITQSRGNSQ- 845
QY 767 TQIGNRLSDNSTATVDNANLNGVNLTSAGSLNNSHFSHQIGDKGTVTYVLEAATWT 826
DB 846 -----VRLTENS-----HMLITGN-----SDVHQDLNNGH----- 871
QY 827 MSQDITLQNLINNSTITLNSAVSSNNTPRRSLETETTTSAHRRNTTLTVNGKLSG 886

DB 872 -----IHLNSA--DNSNNVT-----KXNLTLYN-SLSC 896
QY 887 QGFQFTSLFPGYKSDKILKSNDAEDYLLSVRTGAPETLEQLTVESKONQFSDKL 946
DB 897 NSGFYTLTDLKSGQKDVVTKASNTNFTLQVAKTGEPRH-NELTLPDASKAQ--RDIH 953
QY 947 KETLENDYDAGALRYLVKNDEFRLLNPLKEQELHNDLV----- 987
DB 954 NVSLVNTVDLGAAMYKLRVNGRYDLNYP--EVEKNGQVDTDTNLTTPNNIQLADVSVR 1011
QY 988 -----BARQARTLEAKQVET----- 1004
DB 1012 SNNEIARVDAFPAPPAATPSTETTEIVANSQSKTEVENKEDATETTAQNRVAAE 1071
QY 1005 -----AKTQGE-----PKVSRBA 1020
DB 1072 AKSNVANTQINEVAQSGSEKBTOTPTKETATVEKBAKVBETKQVPAKYSQVSP 1131
QY 1021 RAAPDTLPDS-----LNLAEAK-QAELTAFTQSKAKTKKRSKRAVSPDL 1071
DB 1132 KQSGELVQPAEPARENDPTVNIKEPOSQNTTADTEOPAKETS-----SNVEQPVTE 1185
QY 1072 QSLFLEAALFVIDAPQO-----SEKRLQAEAKQKQK----- 1107
DB 1186 ST--TVNTGNSVVENPANTTPATPTPTNSSESSNKKRNRHSVRSVPNVEPATSSND 1243
QY 1108 -----DLISRYNSALSLSA-----TVNSMLSYODELDELFLV-DQASAVMTNIAD 1154
DB 1244 RSTVALCULTSTNTNAVLSDAKAKQFVALNVGKAVASQHSLSQLENNEGQVNVVNTSM 1303
QY 1155 KERYSDAFRAYQOQNTNLRQIGVQKALANG-RIGAVFESHSDNTPDQVKNHATLTM 1213
DB 1304 NKQISSQYRRFSSKSTQ--TQLGMDQITISNNVGLGVFTYVNSNNEPKATSKN-TLQV 1361
QY 1214 SGFAQY-----QMEDLPQVNVGTGISASKMAEBSKHIRKAINVGVNAYQPRIGQGI 1269
DB 1362 NYSKYVADNMH---VGLDIDGYGFKSQKQTNHAKKPRHTAQGLIAGAKANLQNFPI 1418
QY 1270 QYFVFNVEFTJEBNYSQSEEVKVTPLAFNRYNAGIRVDYTFPTPTNISVKEYPVNY 1329
DB 1419 TPIVAVRSYLSNADBALDQARIKVPISVKTAFQVLDVSTY-HLGEFVPTILSAR- 1476
QY 1330 DYSNANVQTYNLTVLQPRGRYQKVGKAKETLHFOISAFISKSGSQGLGQONVGVK 1389
DB 1477 DNNQSGKLVNNGYDPAYNVENQOYNAGIKLKYHNVKSLIGGLTVAKQAEKQTAELK 1536
QY 1390 LGY 1392
DB 1537 LSF 1539
RESULT 13
US-10-080-505-3
; Sequence 3, Application US/10080505
; Patent No. 6676948
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1541
; TYPE: prt
; ORGANISM: Haemophilus influenzae
US-10-080-505-3

Query Match 23.0%: Score 1673.5; DB 4; Length 1541;
 Best Local Similarity 28.5%: Pred. No. 14e-101;
 Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

6 FRNLPLTACISLIGVSOAMAGHTYFGIDYQYRDPFAENKQPTVGAONIKYVNFQQLV 65
 6 PKCNFLALIVAAALPTTEBALVRDQVQIPDPANKNKPSVGAIVNLVKDNNKDLG 65
 66 TSMITKA-PMIDSVTS-RGVAAVLENQIYVVA-NVQTTDVDRG-AEGN---NEDQH 117
 66 TALPNCIEMIDSVYDVKRIATLIMPQYVGVKNSGVSEIHFQNGNMNNGAKAH 125
 118 RFT-----YKIVKNNY-----KKNLHPYEDDYNHNPRLKHEVTEAIPD-MIS 163
 126 RIVSSENNRYEVEKNEEYPTKANGKVTTEDTQOKREDDYMERLOKTEVVAITEKST 185
 161 NANGSTYSORTYPERVARGSGRFNRNDQKQD-----VAGAYHTL 203
 186 SSDAGTYNDQNKYPAFRLSGSQF---LYKKGNYSLLIHHVEGVGNLKLVDGDAYTG 242
 204 TAGNTNQGAGNGYSTLG-----GDYKAGYGPPLIAGSKGSGSPMFYDAEKQK 257
 243 IAGTPYKVNHNHNGILGFANSKEHSDPKILSODPLTNVAVLGSGSGSPLFYDREKQK 302
 258 LINGILREGNPECKE-NGQVAKSYDPDEIFREDDTSL-----YTRAGNG-VYTI 309
 303 LFIQSYDFAGYNNKSKQWENITYSQPTVDVANKSAGSLIGSKTDYSSNGSTSTTG 362
 310 NDNGQGSITQSGIPESEIKITLANSLPLKEKDVHNPRYDGNITSPRLNGETLYFM 369
 363 GE-----KSLANDLAD-----GKDP-----NHKSYTF-- 386
 370 QKQGSITFASDINOAGGLYFEGNFTVSPNS-NOTWQAGIHVSNSTVTMKVGYEHR 428
 387 EGGSTLJLNNNDIQAGGLFEGDEYVKGTSDDNTTMKAGVYAGKVTWKVHNQYDR 446
 429 LSKIKGTLHYQAKNGKGISVGGKVLTEQADQKQKQFSEIIGVSGRGYQLND 488
 447 LAKIGKTLIVGTGDKKSLKVGQSTVLKQQTNGSG-OHAFASVGVSGSRSLTVLND 505
 489 KQFDIDKYPFGRGRLDNGHSLTFKRIQNTDEGAMIVHNHTTQANVTITGNEIYLP 548
 506 KQVDPMSIYFGRGRLDNGHSLTFKRIQNTDEGAMIVHNHTTQANVTITGNEIYLP 565
 549 N----- 549
 566 NITTPYNDAPDEDNPAFRRIKDGQQLYLNLENTYVALRKAGSTRSELPRKSGSENE 625
 550 -----GNINKLDYRKEIAVNGMFGENDKKNKNGNLNIYKFTTERITLLSG 597
 626 NLYMGKTSDEAKRWVNHINERNGNGVYFG-ERKNGNGLNVTPEKSGSEONRPLTGG 684
 598 GTNLKQDITQTKGLFESGRPTPHAYNHLNKRSEMEGI-----DQGEIWDHDM 647
 685 GTNLKQDITVEKGLFSGRPTPH-----RDLAGISSTKQDHPHANNVVEEDM 736
 648 INRTKRNPOIKGGGAVVS-RVSSISGNTTSSNNAKTFVVPNOQNTITCRSPMTGL 706
 737 INREKATLVNVTGNSLVSGRVANSITSLTASNAQVHIQY-KTGDIVCVRSQYTG 794
 707 TTCQKVDLDTKVLNS-EKTOINGSLINTDACTANYKGLAKNGVNTLTHNSQFTLSNNA 766
 795 VICTDIXLSD-KALNSFNPTNLRGVNLTESANF-VLGGANLFGTIQSRNGSO----- 845
 767 TQIGIRLSDNSTAYVDNANLNGVNLTDASQSLKNSHSHQIQDGKGTVTLENAVWT 826
 846 -----VRLTENS-----HMLTGN-----SDVQDLDLNGH----- 871
 827 MRSDDTLQNTLINSITLNSAYSSANNPRRSLETTTTSABERNTLTVNGKLSG 886
 872 -----IHNLSA-----DNSNNVT-----KXVTLTVN-SLSG 896
 887 QGTQFTSLFGYKSKLKLSDNAGVYLLSVHNTQKEPETLEQTLVESKQNDPLSDKL 946

897 NSSEFYLLTDLNNKQSGKVVVTSATNGFTLQVADKXGEBNH-NEJLTFDASKAQ--RDHL 953
 947 KFTLNDHVDGALRYKLVKNDGEERLHNPKEQELHNLV----- 987
 954 NNSLVGNTVLDGKMKTKLANNNGRIDLYNP-EVEKQNTVDTTNTTTNNIIOADVSPV 1011
 988 -----RAQERTELEAKQVEPT----- 1004
 1012 SNNEIRANDEAPVPPADPATPSETTETVAENKQSKTVEKNEQDATTETAQREVAKE 1071
 1005 -----ATQTCG-----PYRSRRRA 1020
 1072 AASNVKANTQTEVAAGSGSEYETOTTEKETAVEKEKAVETEKTOEPPVTSQVSP 1131
 1021 RAAPDPLPQDS-----LHMLLEAK-QAELTAETQSKYAKTKKRSKRAVSPDPLD 1071
 1132 KQSGSEYTPQAPAREMPNTVNIKEPOSQNTTADTEQPAKETS-----SVEQPYTE 1185
 1072 QSLFALREALVTDAPQ-----SEKDLAQEBARKQKQK----- 1107
 1186 ST-TVNTGNSVVENENTTPATQPTVNSSESNKPKDNRHSVSVPHNVEPATSSND 1243
 1108 -----DLISRSNALSLSA-----TNSMLSYQDELDRLPY-DQASAWMTIAQD 1154
 1244 KSTVALCDLSTVNTNVLSDAPARAQFVALNKGKAVQSHIQLEMMNNGQYNWVNSITSM 1303
 1155 KERYSDAFRAYOQKTNLROIGVQKALANG-RIGAVFESHRSNTPDEQKNAHLLTMM 1213
 1304 KNNYSSQYRRFSKSTQ-TQLGMDQITISNNVQLGSVFTVYRANSNNDKATSKN-TLAV 1361
 1214 SGFAQY-----QMGDLPGVNVGTGISASKRAEBSKRIHRAINYGNASYOFRIGOLGI 1269
 1362 NYSKYVADNMH-----YLGDIGYKGFQSKIQTNHAKAPRHTAQFGJLAKAPNANGGI 1418
 1270 QPYGVNRYEIEKNNQSEEVKVTPLANRNKAGIKVDTFPTPTNISVYEPVNVY 1329
 1419 TPIGVNRYSTLSNADPDLQARIKNPISVKTAFAQVDLSTY-HLGEFVTEPILSARY- 1476
 1330 DVSNANVQTTNLTVYQOPFGRVQKEVGIKAEILLHFOISAFIKSGSQSLGKQOVGVK 1389
 1477 DANGSKIVNVDYDANNVENQOYNAGIKLKYNNVLSLIGLTRYAKABEKQKTAELK 1536
 1390 LGY 1392
 1537 LSF 1539

RESULT 14
 PCT-US95-10661A-3
 : Sequence 3, Application PC/TUS9510661A
 : GENERAL INFORMATION:
 : APPLICANT: Washington University, et al.
 : TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
 : NUMBER OF SEQUENCES: 9
 : CROSS-REFERENCE ADDRESSES:
 : ADDRESSSEE: Flehr, Hobbach, Teat, Albritton & Herbert
 : STREET: 4 Embarcadero Center, Suite 340C
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States
 : ZIP: 94111-4187
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US95/10661A
 : FILING DATE: 16-AUG-1995
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/296,791

FILING DATE: 25-AUG-1994
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Treacartin, Richard F.
 REGISTRATION NUMBER: 31,801
 REFERENCE/DOCKET NUMBER: FP-59941/REF
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 398-3249
 TELEFAX: (415) 781-1989
 TELETYPE: 910 277239
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1541 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 PCT-US95-10651A-3

Query Match: 23.0%; Score 1673.5; DB 5; Length 1541;
 Best Local Similarity: 28.5%; Pred. No. 1,4e-101;
 Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

QY 6 RLNFELTACISLGIVSQAMAGHTFEGIDYQVADPAENKKGKFTGAGNKKVYKQGGQVVG 65
 D6 6 EKLPFALTYAVALTYTTEALVDDVDVQIFDPAENKKGKFSVGAATNYLVQKNNKMDLG 65
 QY 66 TSMTKA-PMIDESVVS-RNGVPAALVENQYIVSAH-NVGYTDVDFG-AEGN-----NPDQ 117
 D6 66 TALNGIPMIDFESVDVDRKIALTLINPYVYGVGVNSGSELHFGKLNKNNNNKAKAH 125
 QY 128 RFT-----KLYKKNY-----KKDLHPEDYDYNPLHKEFTYEAAPID-MTS 160
 D6 128 RDVSEENRIFESVEKEEYPTLNGKVTETEDQKREDDYMPRLDKFVTAAPLEKSTA 165
 QY 161 NMNGSTSDRTKYPERVATGSGQFFMNDQDGDQ-----VAGATYLT 203
 D6 161 SSDAGTYNDQKTPAFATVLSGSGSF---LYKGDNYSLILNNEHVGANNKLVGCPATYTG 242
 QY 204 TASTENQRCAGNGSYLYG-----GPRVKAQVGLPLFAGSKDGSMPFYDAEKOKM 257
 D6 243 IAGTPYKVNHNNGILGFNGSKHEHSDPKGISLSDPLNVAVLGDSGSLFYDREKGM 302
 QY 258 LINGILAEQNPFEQGE-NGFQVAKSYFEIIFERDLTSL-----YTRAGNG-VYTTISG 309
 D6 303 LPLASVDPMAGNKKSKQWQEMKIYKSOFTXDVLNKSAGSLGSKTYVSSNGKSTYTTG 362
 QY 310 NDNGQSGITQSGIPSEIKITILANMSLPKEDKYNHPRYDGNITISPLNNGETLYFMD 369
 D6 363 GE-----KSLNVDLAD-----GDKP-----NKGKSVTF-- 386
 QY 370 QKQSGLIFASDINGAGGLYFESNFTVSPNS-NOTWQAGIHVSNSYTWKXAGVEHDE 428
 D6 387 EGSSTLTIANNNIDQAGGLFPBGDYEVKGTSLNLTWAGAGVSAEKKVTKKYNRPQYDR 446
 QY 429 LSKTGGKTLHQAQKENGKGSISYGDGKVLLEQADDOQKQAFSEIGLVSRGTVQALND 488
 D6 447 LAKTGGKTLVEGTGDNKGSILKVDGIVLTKQNTGSG-QHAFASGIVASRSTLYLND 505
 QY 489 KQPDTKFYFGRGRLDINHSLTFKRLONTDEGAMVYHNTTQAANTTGGESLYLP 548
 D6 506 KQVDPNSIYFGRGRRLDINSLTDPHILNIDDAKLVNHNNTASNTITTGESLITDP 565
 QY 549 N----- 549
 D6 566 NITTPYNDAPDEDNPFAPRIKQGGQLYLNLNMYTYALRKASRSELPLKSGESNEN 625
 QY 550 -----GNINKLDYRKELIYVNGMGEIDKAKHNGSLNLIYFTTEDRTLLSG 597
 D6 626 KLYMGKTSDEARVNVHNNERNGFNQYFGB-EKQKNNGLNLTVFYKSGEQRLFLNG 684
 QY 598 GTNLKGDITQTKGLFSGRPTPHAVNHLNKRMSMGCI-----PGELVMDHDM 647
 D6 685 GTNLKGLTVEKGLTFLSGRPTHA-----FDIAGISSTKQDPHAPREKNEVEVDDW 736

QY 648 INTRFKENFQIKGASAVS-RNVSSTEGNFTVSNANNAFTGVENQNTICTRSMTGL 706
 D6 737 INRFPATMTNTVNASISGRVNAVITSNITASNKAQVHIG--KTGDTYCVASDYTG 754
 QY 707 TTCQKYLDPYVINSIPKQVINGSINTLDNATANVGLAKINGNTLTNHSQTTLSNNA 766
 D6 795 VTCTTKASD-KAANSNPTNLKGVNLTESANP-VLGKANIPLGTIQSRGNSQ----- 845
 QY 767 TQIGTRLSNSTATVFNANINGNVHLNDSQFSLKSHSHQIQGCKATVTLENATWT 826
 D6 846 -----VLTENS-----HMLTGN---SDVHQDLNAGH----- 871
 QY 827 MPDQTLQNLILNASTLTNLSAVSASSNPTRRSS-LETTPTSAHRENTLYVNGKLSG 886
 D6 872 -----IHLNSA--DNSNVVT-----KNYTLTVN-SLGS 896
 QY 887 QGTQFTSLGYSKDKLKLSNDAEGDYLISVRNGKPEPTLEQTLVESKDNPISDLK 946
 D6 897 NGSFYTLTDLSSKQDDVYVTTKSAQNTLQVADKTGEPRH-NELTLPASKAQ--RDHL 953
 QY 947 KFTLENDVDAQALRYKLVDNDEFLNPLIKQELHNDLV----- 987
 D6 954 NVSLVGNATVDLGAWKYKLNVNNGRYOYVNP--EVEKKNQVDTNITTPNNIQADVSVP 1011
 QY 988 -----RAQAKRTLEAKQVEPT----- 1004
 D6 1012 SNNEELARVDAPVPBPAPATSEETETVAENSKQSKTEVKNEDATETTAQNFREAKE 1071
 QY 1005 -----AKTQGE-----PKVRSRAA 1020
 D6 1072 AKSNVANTQINNEAQSGETKETQETKETATVEKKAVENTEKTQEVPRKTSQVSP 1131
 QY 1021 RAAPFDYTPIDS-----LNLAEAK-QAELTAETOKSKATKKVSKRAVPSPLD 1071
 D6 1132 KQOSHTVQPAEPARENDPTVNIKEPOSQNTTADTBQPAKETS-----SNVEQPYE 1185
 QY 1072 QSLPLAEALVIDAPQO-----SEKDLAQEAKQKOK----- 1107
 D6 1186 ST--TVNTGNSVSNPENTTPATTPQTVNSSESNKKNRRSRVSYVPHNVEPATISSND 1243
 QY 1108 -----DLISRYNSALSLSA-----TVNSMLSYODELPLFV-DQASAVNTNIAD 1154
 D6 1244 RSTVALCDLSTNTNNAVLSDAKQAPALANVSKAISOHISOLEMNEGOYVNSNTSM 1303
 QY 1155 KRRYSDAFPAYQOQKNTLRQIGVQKALANG-RIGAVFSHSDNTEDEQKXNATLTMM 1213
 D6 1304 NKRYSSQYRRFSKSTQ--TQGMQDITSNVQLGCVFYVYNSNNDPKATSKG--TLAQY 1361
 QY 1214 SGFAQY-----QMGDLQFGVNVGTGISASKMAEQSKLHKRAINYGVNASYQFRLQGLI 1269
 D6 1362 NFYSKTYADNEM--YLGIDLYGKPOSKLQTNHNAKFPAPHTAOFGLTACKAFNLGNFGI 1418
 QY 1270 QPYFVNVYPIERENYQSEBEVAKTPSLAFNRVYAGIRVDTFTPTDNI-SVRYFVNVY 1329
 D6 1419 PTLVEVRYSLSNADALDQALIKNPLSVKTFAPADYDLSTTY-HLGEFVTPDIISARY 1476
 QY 1330 DYSANAVOTTYNLTVLQDPGRYQKEVGLKAIILAFQISAFISKSGSGLQKQDQVGVK 1389
 D6 1477 DMOQSGKINWGYDPAVNVENQOQYVNAGLKLYHNVKLSLGLVYAKQAEKQATAEIK 1536
 QY 1390 LGY 1392
 D6 1537 LSF 1539

RESULT 15
 US-08-296-791-4
 ; Sequence 4, Application US/08296791;
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene III, Joseph W.
 ; APPLICANT: Fal'kow, Stanley
 ; TITLE OF INVENTION: Haemophilus Adherence and Penetration

TITLE OF INVENTION: Protein
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flehr, Hoback, Test, Albitton & Herbert
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: United States
 ZIP: 94111-4187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/296,791
 FILING DATE: 25-AUG-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Treccartin, Richard F.
 REGISTRATION NUMBER: 31,801
 REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1545 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 US-08-296-791-4

Query Match 22.9%; Score 1662.5; DB 3; Length 1545;
 Best Local Similarity 28.6%; Pred. No. 7.7e-101;
 Matches 462; Conservative 246; Mismatches 511; Indels 447; Gaps 59;

6 ELNPLFACISLIGVQAMAGHTVEGIDYQYVDEFAENKGETVGNQIKYNNKQGLVNG 65
 6 FKNPLALTYAALTYTEALVDDVDYVITFDEFAENKGFESVGAINEVADKNNRILG 65
 66 TSMTKA-PMIDESVVS-RNGVVALVENQYIVSAH-NGVYDVFQ-ASGN---NPDQH 117
 66 NVLPNG-PMIDESVWVDKRIATLVNPOYVGVKVSNGSELHFNGLNKNNNKAKAH 125
 118 RFT-----YKIKKNNY-----KKNLHFEEDYHNPRLKHYTAAPTD-MTS 160
 126 KDVSSENNRYTYEKREPYTLNKGAVTTEDQAQRREDDYMWRLKPFTEVAPTEASTD 185
 161 NMNGSTYSRSTKPYRVRIGSRQF-----WRNDQ-DKGDQVAGAYHVLANTH 209
 166 SSTAIGYNNNDKYPFYVALGSGTQFIYENGTYELMLGKQKSD--AGGIVLKLGNAY 243
 210 NQAGANGVSY-----LGGDVRAKGEY-GPLPIAGSK-----GDSGSPMFLYDAEK 254
 244 TYGIAGTPYEVVHENDGLIFGNSNNEYLNPKELISKKPLTNVAVALGDSGSPFLYDREK 303
 255 CKWJLNGILREBNPEKGE-NGFOLVAKSYFPEIPEEDLTSL-----YTRAGNG-VTT 306
 304 GKMLFEGSYDYVAGNKKSKWQEMNLYKPEFAKITYEYSAGSLGSKTDYSNSGKTST 363
 307 ISGNDNGGSLTGKSGISEIKITLNMSLPLKEKCYHNPBRIDGPNISPLNNGETLV 366
 364 ITGGE-----KSLNPLDLD-----GDKRP-----NHGKSVT 389
 367 FMDQKGSILFASDINQAGELYFEGNFTVSNS-NQTOAGIHVENSITVTKVNGVE 425
 390 F-EGSGTLTNLNNIDQAGELFEEDYEVKTSNTTMMGAGVSAABOKTVTKVANNPQ 447
 426 HDRLSIKGKGLHYVAKGEEKKSGISYGDQKVLTDQADDQANKQAFSELGYSRGVVOU 485
 448 YDLAKIGKGLIYVGGTGDNGSLKVGDSGTILAKQNTGSGG-CHAFASVGIIVSGRSTLV 506

486 NDDKQFTDKEFYRGRGRDLNLSHSLTEFKRIQNTDEBAMTVNNTQOANVTITNGESI 545
 507 NDDKQVPSNIVYFGRGRDLNLSHSLTEFKRIQNTDEBAMTVNNTQOANVTITNGESI 566
 546 VLPNG----- 550
 567 TDPNNSVITYVPLEDDNPVARIQIKYQOLYFNEENRTYALAKKQASIRSEFPQNRGES 626
 551 -----NNINLDYRKEIAVNGMPGETDKNKNRNLNLYXPTEDRTLL 594
 627 NNSLWNGTEKQADQAKNMMHNNERNNGPNQYFE-EGGNGNINLVFKKSGKQNF 685
 535 LSGGTNLKSDITQYTKGKLPFSGRPTPHANNLKNKWSMEGI-----POGEIYWD 644
 686 LVGGTINMGLWVQGTFLFSGRPTPHA-----RDIAGISSTKDKSHFSENNVEVE 737
 645 HDWINTPEKRNPOKGSAAVS-RVSSISEGWNVTYSNANATFPGVVPQOVTITRSDW 703
 738 DDMWNRFPATNIVNTNATLYSGRVESITSNITASNAKVAHIGY-YAGDITVCVRSDY 795
 704 TGLTTCQKVDLDTVYVINSLEPTQINGSINTLDNATANYGLAKLNGVTLTNHSGFTLS 763
 796 TGYVTCITDKSD-KALNSPPTNLGANNLTESANF-VLGKANLGTITQSGKNSQ--- 849
 764 NNATQIGTRLDNSTATVDNANLNGNVEHTDSAQFSLKNSHSHQIQGDKGTVTLEMA 823
 850 -----VALTEMS-----HMLTGN--SDVEQLDLANGH----- 875
 824 TWMTSPDITLQVLILNNSITLNSAASANNTPRRSLTETTPPSAHRNTLTVNGK 883
 876 -----HLNSA--DQSNVNT-----KXNLTVN-S 897
 884 LSGQGTFOFTSLFQYKSDYLKLSNDAEGYLLSVRTGKEPETTELEOTLYESKDNQPLS 943
 898 LSGNGSFYTLDLDSKQGDVAVVTKSAIGNFTLVQANXTEPERH-NELTLDPAKQ--R 954
 944 DKLFTLNDYVDGALRYLQVNDGEFRLNPIKQGEHLNDIV----- 987
 955 DHALNSLVGNTVDLGAKKYDLRVNNGRYDLNVP--EYKKNQVDTDTNITPPNNIQADVP 1012
 988 -----RAQAKRTLEKQVEPI----- 1004
 1013 SVPSNNEIARUDAPFPAPATSETTEVEANSKQSKTEVNEKQDATTETTAQNRREV 1072
 1005 -----AKTQGE-----PKYRSR 1017
 1073 AKKAKSNVANTQINBAVAGSGSETKETQTEKTLATVEKBAKAVETKQVEPKTQ 1132
 1018 RABARAFDTLPDS-----LNLALAK-QALTAFTQSKAKTYKVSRRANVESPD 1068
 1133 VSPKQSGETVQPAEPARENDPTVNIKEPOSQNTNTADTEQPAKETS-----SNVEOP 1186
 1069 LLDQSLFALAEALVEIDAPOQ-----SEKDRLAQEEAEKQKXK----- 1107
 1187 VTEST--TYWNGNVANPENPTTPTPTPNSSSKKRNKRHSVSVNHNVEPATTS 1244
 1108 -----DLISRYNSALSELISA-----TYNSMLSVODELRLFV-DQASAVWTNI 1151
 1245 SDRSTVALCDLITSTNTNVLSDAKAKQFVALNVGAVSGHISGLBNNNGQVNNWVEN 1304
 1152 AADKRRYSDAFRAYQOQKTNLRQIGVQKALANG-RIGAVPSHSSNDTPEVQKHAATL 1210
 1305 TSNKNNYSSQYRRFSSKSTQ--TQIGMDQTTISNNVQGLGVFTYVNSNNKPKATSKN-TL 1362
 1211 TWMSGFAQY-----QWEDLPQVNVGTGISAKMAEQSKIRKAKLVNGNVASYOFRLO 1266
 1363 AQVNYYSKYVADNHN---YLGIDLGKGFQSKQTMHNAKPKRHHAPGLTAKAKANLON 1419
 1267 LGIQYFPVNRYPTEBNTQSEEVVKTPESLAFNRNAGIRVDYTFPTNTNISVKEYPFV 1326
 1420 FGITPEIVGVRSYLSNADPALDQAFIKVNPISVKTAFQVDSLSTY-HLGEFVPTIILSA 1478
 1327 NYVDVSANANQTTVNLTVLQGFQGRVWQXGVAKELIHFQISAFISKSGSGLSKQQNV 1386

Db	1479	RY-DANQSGSKINWNGYDFAYVVENOOYNAGIKLKHYHVKLSJGGJLTAKQAEKOKTA	1537
Qy	1387	GYKLGX	1392
Db	1538	EIKLSF	1543

Search completed: June 23, 2004, 17:30:38
Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw mode)

Run on: June 23, 2004, 17:31:53 ; Search time 60 Seconds
(without alignments)
6559.087 Million cell updates/sec

Title: US-10-645-655-2

Sequence: 1 MKKTVRLNPLTACTISIV.....SQSGQKQGVKGLGYM 1394

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1-63542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Database :

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Published Applications AA:*

1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/prodata/1/pubpaa/PTC_NEW_PUB.pep:*

3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/prodata/1/pubpaa/PTCUS_PUBCOMB.pep:*

7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/prodata/1/pubpaa/US09D_PUBCOMB.pep:*

13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pep:*

17: /cgn2_6/prodata/1/pubpaa/US10E_PUBCOMB.pep:*

18: /cgn2_6/prodata/1/pubpaa/US10F_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7273	100.0	1394	10	US-09-839-996-2
2	7273	100.0	1394	12	US-10-645-655-2
3	7273	100.0	1394	14	US-10-080-505-2
4	7246.5	99.6	1395	14	US-10-080-505-7
5	6115	84.1	1434	14	US-10-080-505-9
6	5827.5	83.1	1411	14	US-10-080-505-17
7	5807	79.8	1395	14	US-10-080-505-13
8	5303.5	72.9	1395	14	US-10-080-505-15
9	5297.5	72.8	1395	14	US-10-080-505-11
10	4178	57.4	1449	12	US-10-282-122A-65711
11	4085	56.2	1457	16	US-10-220-481-71
12	4072.5	56.0	1458	12	US-10-282-122A-65027
13	4044.5	55.6	1439	16	US-10-220-481-73
14	1673.5	23.0	1541	10	US-09-839-996-3
15	1673.5	23.0	1541	12	US-10-645-655-3

16	1673.5	23.0	1541	14	US-10-080-505-3	Sequence 3, Appli
17	1662.5	22.9	1545	10	US-09-839-996-4	Sequence 4, Appli
18	1662.5	22.9	1545	12	US-10-645-655-4	Sequence 4, Appli
19	1662.5	22.9	1545	14	US-10-080-505-4	Sequence 4, Appli
20	1615.5	22.2	1593	12	US-10-282-122A-65262	Sequence 6, Appli
21	1609	22.1	1694	12	US-10-282-122A-58377	Sequence 58377, A
22	1605	22.1	1702	10	US-09-839-996-5	Sequence 5, Appli
23	1605	22.1	1702	12	US-10-645-655-5	Sequence 5, Appli
24	1605	22.1	1702	14	US-10-080-505-5	Sequence 5, Appli
25	1551	21.3	1848	10	US-09-839-996-6	Sequence 6, Appli
26	1551	21.3	1848	12	US-10-645-655-6	Sequence 6, Appli
27	1551	21.3	1848	14	US-10-080-505-6	Sequence 6, Appli
28	1536.5	21.1	1773	12	US-10-282-122A-65817	Sequence 65817, A
29	1089	15.0	1371	14	US-10-238-075-891	Sequence 891, App
30	1051	14.5	1376	14	US-10-238-075-560	Sequence 560, App
31	346.5	4.8	73	14	US-10-080-505-44	Sequence 44, Appl
32	344.5	4.7	1569	12	US-10-282-122A-43126	Sequence 43126, A
33	344.5	4.7	1569	14	US-10-287-274-312	Sequence 312, App
34	326.5	4.5	1430	12	US-10-282-122A-78451	Sequence 78451, A
35	292.5	4.0	3795	12	US-10-282-122A-77944	Sequence 77944, A
36	288	4.0	836	12	US-10-282-122A-56613	Sequence 56613, A
37	286.5	3.9	978	12	US-10-282-122A-75356	Sequence 75356, A
38	286.5	3.9	1536	13	US-10-092-880-2	Sequence 2, Appli
39	285.5	3.9	1004	14	US-10-193-764-53	Sequence 53, Appli
40	285.5	3.9	1010	14	US-10-193-764-51	Sequence 51, Appli
41	283.5	3.9	1222	14	US-10-193-764-37	Sequence 37, Appli
42	283.5	3.9	1228	14	US-10-193-764-34	Sequence 34, Appli
43	283.5	3.9	1536	13	US-10-193-764-63	Sequence 63, Appli
44	283	3.9	1477	14	US-10-092-880-4	Sequence 4, Appli
45	280	3.8	1477	14	US-10-193-764-67	Sequence 67, Appli

ALIGNMENTS

RESULT 1
US-09-839-996-2
Sequence 2, Application US/09839996
Publication No. US2003009010A1
GENERAL INFORMATION:
APPLICANT: St. Gene III, Joseph W.
Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fleury, Heibach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,996
FILING DATE: 20-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:

US 1064565509P1



Creation date: 07-15-2004
Indexing Officer: SMOHAMMED - SUAD MOHAMMED
Team: 1600PrintWorkingFolder
Dossier: 10645655

Legal Date: 07-07-2004

No.	Doccoder	Number of pages
1	SRNT	1

Total number of pages: 1

Remarks:

Order of re-scan issued on